

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 March 2003 (20.03.2003)

PCT

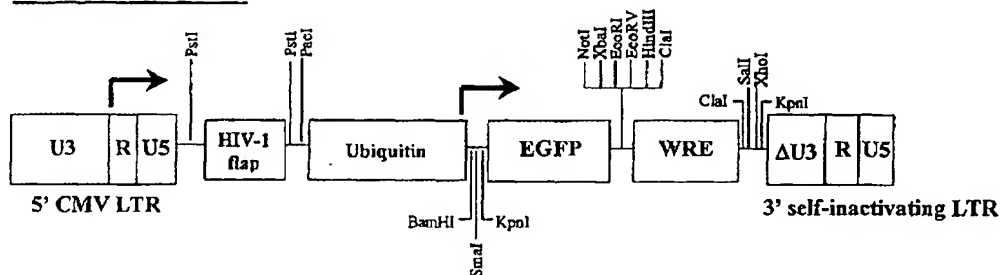
(10) International Publication Number  
**WO 03/022040 A2**

- (51) International Patent Classification<sup>7</sup>: **A01K** [ES/US]; 1821 1/2 Wayne Avenue, South Pasadena, CA 91030 (US). **PEASE, Shirley** [GB/US]; 341 Hill Street, Monrovia, CA 91016 (US).
- (21) International Application Number: PCT/US02/29130
- (22) International Filing Date: 12 September 2002 (12.09.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/322,031 13 September 2001 (13.09.2001) US  
60/347,782 9 January 2002 (09.01.2002) US
- (71) Applicant (for all designated States except US): **CALIFORNIA INSTITUTE OF TECHNOLOGY** [US/US]; 1200 E. California Boulevard, Pasadena, CA 91125 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **BALTIMORE, David** [US/US]; 415 S. Hill Ave., Pasadena, CA 91106 (US). **HONG, Elizabeth, J.** [US/US]; 30 Barrington Drive, Bridgewater, NJ 08807 (US). **LOIS-CABALLE, Carlos**
- (74) Agent: **ALTMAN, Daniel, E.**; Knobbe, Martens, Olson and Bear, LLP, 2040 Main Street, Fourteenth Floor, Irvine, CA 92614 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

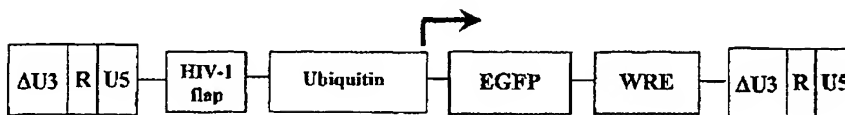
[Continued on next page]

(54) Title: METHOD FOR PRODUCING TRANSGENIC ANIMALS

Viral vector: FUGW



Integrated proviral locus



Note: Diagrams are not to scale.

(57) Abstract: The present invention relates to methods for producing transgenic animals using retroviral constructs engineered to carry a transgene(s) of interest.



WO 03/022040 A2



**Declaration under Rule 4.17:**

- *of inventorship (Rule 4.17(iv)) for US only*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

- *without international search report and to be republished upon receipt of that report*

## METHOD FOR PRODUCING TRANSGENIC ANIMALS

5

### Statement of Government Support

This invention was made with government support under Grant Number GM39458 awarded by the National Institutes of Health. The United States Government has certain rights in the invention.

10

### Background of the Invention

#### Field of the Invention

The present invention relates to methods for generating transgenic animals using viral constructs engineered to carry the transgene(s) of interest.

15

#### Description of the Related Art

Early transgenic experiments used an oncoretrovirus to introduce the gene of interest into embryonic cells (Jaenisch Proc. Natl. Acad. Sci. USA 73:1260-1264 (1976)). In a typical experiment an engineered Moloney strain of mouse leukemia virus (MoMLV) was injected into the blastocyst cavity of mice. While the transgene was often integrated into the genome of the resulting mice, no gene expression could be detected.

20

Today, the majority of transgenic animals are made using direct injection technology (Gordon and Ruddle Science 214:1244-1246 (1981)). Briefly, a DNA construct carrying the gene of interest is injected directly into the pronucleus of a single-cell zygote. The cell is then implanted into a pseudo-pregnant female and the resulting progeny is analyzed for expression of the gene.

25

While this method achieves both integration and expression of the transgene, there are a number of significant drawbacks to the direct injection technique. First, in order to carry out the technique it is necessary to inject

30

DNA directly into the pronucleus. This is possible in some specific strains of mice, such as Black6 x BDA, because the male pronucleus is visible. However, in other animals and other strains of mice the pronucleus is less visible, making the technique extremely difficult. Further, the injection  
5 requires the assistance of a skilled technician and a significant investment in equipment; micromanipulators are necessary to hold the cell and the injection pipette and a pressure source is required that can deliver picoliter amounts of DNA solution.

A second equally significant problem with the direct injection method  
10 is the low percentage of injected zygotes that produce transgenic animals. The injection pipette must go through the zona pellucida, the cell membrane and the nuclear envelope. Thus only 80-90% of mouse cells survive the injection. Other animal cells are less hardy and the survival rates are somewhat lower, with about 60% survival for rats and 40-50% for cows. In mice, of the  
15 original zygotes, about 25% are successfully injected and implanted in a pseudopregnant female. About 20% of the resulting animals have the transgene integrated into their genome. Of these, about 20% will express the gene. However, even if the animals express the gene, it is possible that the expression pattern will not be useful. Thus, only about 1% of injected zygotes  
20 result in transgenic animals that express the gene of interest. This low efficiency of transgenesis is particularly troubling for larger animals, such as pigs, cows or goats, in which obtaining large numbers of embryos is not possible (see, e.g., Wall et al. J. Cell. Biochem. 49:113 (1992)).

In addition, direct pronuclear injection is not possible for many types  
25 of animals, including birds.

#### Summary of the Invention

One aspect of the present invention concerns a method of producing a transgenic animal. In the preferred embodiment a packaging cell line is  
30 transfected with a retroviral construct, recombinant retrovirus is collected, and an embryonic cell is infected with the recombinant retrovirus. The retroviral construct preferably comprises the R and U5 sequences from a 5' lentiviral

long terminal repeat (LTR) and a self-inactivating lentiviral 3' LTR. Further, the self-inactivating 3' LTR preferably comprises a U3 element with a deletion of its enhancer sequence. In one embodiment the LTR sequences are from HIV.

5           In one embodiment the retroviral construct comprises a gene that is desirably expressed in the transgenic animal. In this embodiment the retroviral construct may also comprise an internal promoter and/or enhancer. In one embodiment the internal promoter is ubiquitous. The ubiquitous promoter may be selected from the group consisting of the ubiquitin promoter, the CMV  $\beta$ -actin promoter and the pgk promoter. In another embodiment the  
10           internal promoter is tissue-specific. The tissue specific promoter may be selected from the group consisting of the lck promoter, the myogenin promoter and the thyl promoter.

          In addition, the recombinant retrovirus may be pseudotyped. Thus, in  
15           one embodiment the recombinant retrovirus is pseudotyped with the vesicular stomatitis virus envelope glycoprotein. In another embodiment the recombinant retrovirus is pseudotyped with a mutant ecotropic envelope protein, preferably ecotropic envelope protein 4.17.

          The viral construct may comprise one or more additional genetic  
20           elements. In one embodiment the viral construct comprises a promoter operably linked to the R and U5 5' LTR sequences, preferably a CMV promoter sequence. An enhancer, preferably a CMV enhancer sequence, may also be included in the viral construct.

          In another embodiment the viral construct comprises a woodchuck  
25           hepatitis virus enhancer element sequence. In yet another embodiment the viral construct comprises a tRNA amber suppressor sequence.

          The viral construct may additionally comprise a reporter gene operably linked to the internal promoter. The reporter gene may encode be a fluorescent protein, preferably green fluorescent protein.

30           In one embodiment the embryonic cell is infected by injecting the recombinant retrovirus between the zona pellucida and the cell membrane. In another embodiment the embryonic cell is infected by removing the zona

pellucida, preferably be enzymatic digestion, and incubating the denuded cell in a solution containing the recombinant retrovirus.

5 In another aspect the invention concerns a method of producing a transgenic mammal comprising removing the zona pellucida from an embryonic cell, contacting the embryonic cell with a modified retrovirus and implanting the embryonic cell in a pseudopregnant female. The embryonic cell is preferably a one-cell embryo. The retrovirus is preferably a modified lentivirus. The modified lentivirus is preferably produced by transfecting a packaging cell line with a viral construct. The viral construct may comprise  
10 the R and U5 sequences from a lentiviral 5' LTR, an internal promoter, a gene of interest and a self inactivating lentiviral 3' LTR.

In another aspect the invention concerns a method of producing a transgenic animal comprising injecting a modified lentivirus into the perivitelline space of an embryonic cell.

15 In yet another aspect the invention concerns a transgenic animal made by any of the disclosed methods. Thus the transgenic animal preferably has a genome that comprises proviral DNA. The proviral DNA may comprise a self-inactivating lentiviral 3' LTR, such as a self-inactivating HIV 3' LTR. In particular, the self-inactivating 3' LTR may have a deletion of its enhancer  
20 sequence.

In a further aspect the invention concerns a viral construct comprising CMV enhancer/promoter sequences, the R and U5 sequences from the 5' HIV LTR, the HIV'1 flap signal, an internal enhancer and/or promoter, an exogenous gene of interest, the woodchuck hepatitis virus responsive element,  
25 a tRNA amber suppressor sequence, a U3 element with a deletion of its enhancer sequence, the chicken  $\beta$ -globin insulator and the R and U5 sequences of the 3' HIV LTR.

In still further aspects the invention concerns methods of making transgenic birds, fish and other animals. In one embodiment a transgenic bird  
30 is produced by a method comprising transfecting a packaging cell line with a viral construct, recovering recombinant retroviral particles, and infecting a bird egg with the recombinant retroviral particles. Preferably the viral

construct comprises the R and U5 sequences from a 5' lentiviral LTR and a self-inactivating 3' lentiviral LTR. Infecting the bird egg preferably comprises contacting the embryonic blastodisc of the bird egg with the retroviral particles.

5           In another embodiment a transgenic fish is produced by a method that comprises transfecting a packaging cell line with a viral construct, recovering recombinant retroviral particles, and infecting a fish egg with the recombinant retroviral particles. Preferably the viral construct comprises the R and U5 sequences from a 5' lentiviral LTR and a self-inactivating 3' lentiviral LTR.  
10       Infecting the fish egg preferably comprises delivering the retroviral particles to the space between the chorion and the cell membrane of the fish egg.

          In another aspect the invention concerns methods of isolating genes expressed in particular tissues. In one embodiment embryonic cells are infected with a modified retrovirus whose genome comprises a self-  
15       inactivating lentiviral 3' LTR, a splice acceptor sequence and a reporter gene. Tissues from the developing embryo or mature mammal that express the reporter gene are identified and the gene sequences that flank the provirus are identified.

          In a further aspect the invention concerns methods of identifying genes  
20       that play a role in biological processes. In one embodiment embryonic cells are infected with a modified retrovirus whose genome comprises a self-inactivating lentiviral 3' LTR. A phenotype in the resulting transgenic animal is observed and the gene that was disrupted by the provirus is identified by sequencing the nucleic acid flanking the provirus.

25

#### Brief Description of the Drawings

Figure 1A is a diagram of the FUGW viral construct. Figure 1B is a diagram of the provirus that is integrated into the host genome after infection with recombinant virus prepared with the FUGW viral construct of Figure 1A.

30       Figure 2 is a Southern blot analysis of proviral transgene insertions in the founder generation of mice generated by injecting recombinant lentivirus into the perivitelline space of one-cell embryos. Genomic DNA from each

animal was digested with either PstI (left) or BamHI (right), and probed with a GFP+WRE sequence. All PstI and BamHI sites in the provirus are located 5' to the GFP gene. Plus signs above each lane indicate GFP expression in the animal detectable by viewing under conventional epifluorescence.

5           Figure 3 is a Southern blot analysis of proviral transgene insertions in the founder generation of a second group of mice generated by injecting recombinant lentivirus into the perivitelline space of one-cell embryos. Genomic DNA from each animal was digested with BamHI and probed with a GFP+WRE sequence. All BamHI sites in the provirus are located 5' to the  
10       GFP gene. Plus signs above each lane indicate GFP expression in the animal detectable by viewing under conventional fluorescence. Of the 56 founder animals in this experiment, 45 or 80.4% are transgenic. Of these 45 transgenic animals, 41 or 91.1% express GFP at detectable levels. Lanes marked "C" are positive plasmid controls.

15           Figure 4 shows ubiquitous GFP expression in rats derived from the delivery of modified lentivirus to single-cell embryos *in vitro*. Figure 4A shows brightfield (BF) and fluorescent images of the paws of newborn rats derived from a FUGW-injected embryo. Pup R4, carrying 4 copies of the proviral insert, expresses GFP in the paw, as well as all other tissues and  
20       organs examined. A littermate (R3) carrying no transgene is included for comparison. Figure 4B shows a Southern blot analysis of proviral insertions in rats generated by injection of FUGW lentivirus into the perivitelline space of single-cell embryos. Genomic DNA was digested with PstI and hybridized with a GFP+WRE probe. Plus signs above each lane indicate GFP expression  
25       in the animal detectable by direct viewing under a fluorescent microscope.

          Figure 5 is a Southern blot analysis of proviral transgene insertions in the founder generation of mice generated by incubating denuded embryos in media comprising recombinant lentivirus produced with the FUGW viral construct. Genomic DNA from each animal was digested with PstI and  
30       probed with a GFP+WRE sequence. All PstI sites in the provirus are located 5' to the GFP gene. The ratios above the lanes indicate the dilution of the virus from  $1 \times 10^6$  pfu/ $\mu$ l.



Figure 6 shows GFP expression in major tissues and organs of a founder mouse. The mouse was perfused intracardially with PBS, pH 7.4, and then 3% paraformaldehyde, and viewed immediately under a fluorescent dissecting microscope. The particular mouse shown was generated by co-incubation of the denuded embryo with the lentiviral suspension and contains 8 proviral insertions. A wildtype animal, identically perfused and photographed is included for comparison.

Figure 7 is a Southern blot analysis of proviral transgene insertions in the F1 progeny of founder transgenic mice, showing that the F1 progeny inherit the proviral transgene in a Mendelian fashion. The founder mice were generated by injection of FUGW lentivirus into the perivitelline space of single-cell embryos. Genomic DNA from each animal was digested with BamHI and probed with GFP+WRE sequence. All BamHI sites are located 5' to the GFP gene. The first numbered lane in each group is the P0 founder animal, while the lettered lanes represent progeny resulting from outcrossing that founder animal to a wildtype animal. Plus signs above each lane indicate GFP expression in that animal detectable by direct viewing of the live animal under a conventional epifluorescence microscope.

Figure 8 shows that transgenic mice give rise to transgenic progeny that express the transgene. This indicates that the transgene can go through an entire round of gametogenesis and development without being silenced. Expression of the transgene was determined based on GFP expression in the newborn pup. The pup imaged here is descended from an animal with 10 proviral insertions.

Figure 9 is a Southern blot analysis of proviral transgene insertions in the founder generation of mice generated using a lentiviral vector containing the myogenin promoter driving a histone2B-GFP fusion. Embryos were recovered from the uterus at embryonic day 11.5 ("E11.5"). The litter consisted of 6 animals, all of which were transgenic. Genomic DNA from each animal was digested with BamHI and probed with a GFP+WRE sequence. A BamHI site is located within the histone2B gene, 5' of the GFP sequence in the provirus. Plus signs above the lanes indicate positive plasmid

controls. Animals 5 and 6 were positive for tissue-specific GFP expression at embryonic day 11.5 when viewed as a whole mount under an inverted fluorescent microscope, and animal 5 expressed more highly than animal 6.

5 Figure 10 shows the GFP expression pattern in an E11.5 mouse embryo derived from the perivitelline space injection of lentivirus carrying a histone2B-GFP fusion construct under the control of the myogenin promoter (Yee et al. *Genes and Dev.* 6:1277-1289 (1993)). GFP expression is localized to the somites and can be seen in the emerging muscles in the limb buds, eye and jaw.

10 Figure 11 shows immunofluorescence with an antibody against GFP in a cross-section through an E11.5 embryo carrying the myogenin promoter driving GFP expression. Embryos were derived from single-cell zygotes injected with recombinant lentivirus in the perivitelline space. Embryos were fixed in 3% paraformaldehyde, cryoprotected in 30% sucrose overnight and 30  $\mu$ m sections were cut on a cryostat. Sections were incubated with a polyclonal antibody against GFP and probed with  $\alpha$ -rabbit secondary antibody conjugated to a rhodamine fluorophore. Images on the left are sections as viewed under a rhodamine filter, while images on the right show the nuclear counterstain Hoechst-33342 for each corresponding section. The animal carried 6 proviral  
15 insertions of the myogenin-GFP construct. Specific staining of somite tissues can be seen, with the exclusion of the stain from flanking skin and bone tissues.

Figure 12 also shows immunofluorescence with an antibody against GFP in cross sections of an E11.5 embryo carrying a myogenin promoter driving H2B-GFP. Lack of staining in the viscera is noteworthy.  
25

Figure 13 also shows immunofluorescence with an antibody against GFP in cross sections of an E11.5 embryo carrying a myogenin promoter driving H2B-GFP. Specific staining of somites on either side of the neural tube can be visualized.

30 Figure 14 shows H2B-GFP expression in the extraembryonic tissue of developing zebra finch.

Figure 15 shows H2B-GFP expression inside of the zebra finch embryo.

Figure 16 shows the nucleotide sequence of GFP.

Figure 17 shows the nucleotide sequence of H2B-GFP.

5        Figure 18A shows the nucleotide sequence of HIV NL4.3 flap and 18B shows the nucleotide sequence of WRE.

Figure 19A shows the nucleotide sequence of the myogenin promoter and 19B shows a partial nucleotide sequence of the Lck promoter.

10       Figure 20 shows the nucleotide sequence of the human ubiquitin promoter.

Figure 21 shows the nucleotide sequence of the HIV-1 flap + ubiquitin + GFP + WRE construct.

#### Detailed Description of the Preferred Embodiment

15        Retroviruses are enveloped RNA viruses that are capable of infecting animal cells. When a retrovirus infects a cell, its RNA genome is converted into a double-stranded linear DNA form by reverse transcription. The DNA form is integrated into the host cell genome as a provirus. The present invention is based on the discovery that recombinant retroviruses can be used  
20       to create transgenic animals. Transgenic animals resulting from the methods of the present invention have one or more copies of the transgene of interest integrated in their genome.

Previous transgenic technology is not commercially practical in larger animals, such as monkeys, dogs, poultry, cows, pigs or sheep. Furthermore,  
25       previous transgenic methods are not applicable to poultry. Thus, the methods of the present invention will find great commercial application, for example in biotechnology and agriculture. The present methods may be used to introduce the gene of choice into animals in order to confer upon them desired attributes. For example, the described methods may be used to confer disease resistance.  
30       In biotechnology, the ability to rapidly develop large numbers of transgenic animals, particular higher order animals such as monkeys, will allow for the analysis of gene function and the evaluation of compounds that potentially

modulate gene expression, protein function, or are useful in treating a disease or disorder. Two types of assays in which the methods of the present invention are particularly useful are gene trap assays and large-scale mutagenesis screens, each of which is described below.

5

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Any methods, devices and materials similar or equivalent to those described herein can be used in the practice of this invention.

10

By "transgene" is meant any nucleotide or DNA sequence that is integrated into one or more chromosomes of a host cell by human intervention, such as by the methods of the present invention. In one embodiment the transgene comprises a "gene of interest." A "gene of interest" is a nucleic acid sequence that encodes a protein or other molecule that is desirable for integration and/or expression in a host cell. In this embodiment the gene of interest is generally operatively linked to other sequences that are useful for obtaining the desired expression of the gene of interest, such as transcriptional regulatory sequences. In another embodiment the transgene can be a DNA sequence that is used to mark the chromosome where it has integrated. In this situation, the transgene does not have to comprise a gene that encodes a protein that can be expressed. This use of the transgene as a molecular tag has numerous applications, for example for mutagenesis studies as described below.

15

20

25

30

The term "transgenic" is used herein to describe the property of harboring a transgene. For instance, a "transgenic organism" is any animal, including mammals, fish, birds and amphibians, in which one or more of the cells of the animal contain nucleic acid introduced by way of human intervention, such as by the methods described herein. In the typical transgenic animal, the transgene causes the cell to express or overexpress a recombinant protein. However for some applications, such as the mutagenesis

studies described below, it is not necessary or desirable for the transgenic organism to express a recombinant protein.

The terms "founder," "founder animal" and "founder line" refer to those animals that are mature products of the embryos or oocytes to which the transgene was added, i.e. those animals that grew from the embryos or oocytes into which DNA was inserted.

The terms "progeny" and "progeny of the transgenic animal" refer to any and all offspring of every generation subsequent to the originally transformed animal.

The term "animal" is used in its broadest sense and refers to all animals including mammals, birds, fish, reptiles and amphibians.

The term "mammal" refers to all members of the class Mammalia and includes any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports or pet animals, such as mouse, rabbit, pig, sheep, goat, cattle and higher primates.

The term "oocyte" refers to a female gamete cell and includes primary oocytes, secondary oocytes and mature, unfertilized ovum. As used herein, the term "egg" when used in reference to a mammalian egg, means an oocyte surrounded by a zona pellucida. The term "zygote" refers to a fertilized ovum. The term "embryo" broadly refers to an animal in the early stages of development.

"Perivitelline space" refers to the space located between the zona pellucida and the cell membrane of a mammalian egg or embryonic cell.

"Target cell" or "host cell" means a cell that is to be transformed using the methods and compositions of the invention.

"Lentivirus" refers to a genus of retroviruses that are capable of infecting dividing and non-dividing cells. Several examples of lentiviruses include HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2), the etiologic agent of the human acquired immunodeficiency syndrome (AIDS); visna-maedi, which causes encephalitis (visna) or pneumonia (maedi) in sheep, the caprine arthritis-encephalitis virus, which causes immune deficiency, arthritis, and encephalopathy in goats; equine

infectious anemia virus, which causes autoimmune hemolytic anemia, and encephalopathy in horses; feline immunodeficiency virus (FIV), which causes immune deficiency in cats; bovine immune deficiency virus (BIV), which causes lymphadenopathy, lymphocytosis, and possibly central nervous system infection in cattle; and simian immunodeficiency virus (SIV), which cause  
5 immune deficiency and encephalopathy in sub-human primates.

A lentiviral genome is generally organized into a 5' long terminal repeat (LTR), the gag gene, the pol gene, the env gene, the accessory genes (nef, vif, vpr, vpu) and a 3' LTR. The viral LTR is divided into three regions  
10 called U3, R and U5. The U3 region contains the enhancer and promoter elements. The U5 region contains the polyadenylation signals. The R (repeat) region separates the U3 and U5 regions and transcribed sequences of the R region appear at both the 5' and 3' ends of the viral RNA. See, for example, "RNA Viruses: A Practical Approach" (Alan J. Cann, Ed., Oxford University Press, (2000)), O Narayan and Clements J. Gen. Virology 70:1617-1639  
15 (1989), Fields et al. Fundamental Virology Raven Press. (1990), Miyoshi H, Blomer U, Takahashi M, Gage FH, Verma IM. *J Virol.* 72(10):8150-7 (1998), U.S. Patent No. 6,013,516.

"Virion," "viral particle" and "retroviral particle" are used herein to refer to a single virus comprising an RNA genome, *pol* gene derived proteins, *gag* gene derived proteins and a lipid bilayer displaying an envelope (glyco)protein. The RNA genome is usually a recombinant RNA genome and thus may contain an RNA sequence that is exogenous to the native viral genome. The RNA genome may also comprise a defective endogenous viral  
25 sequence.

A "pseudotyped" retrovirus is a retroviral particle having an envelope protein that is from a virus other than the virus from which the RNA genome is derived. The envelope protein may be from a different retrovirus or from a non-retroviral virus. A preferred envelope protein is the vesicular stomatitis virus G (VSV G) protein. However, to eliminate the possibility of human infection, viruses can alternatively be pseudotyped with ecotropic envelope protein that limit infection to a specific species, such as mice or birds.  
30

The term "provirus" is used to refer to a duplex DNA sequence present in a eukaryotic chromosome that corresponds to the genome of an RNA retrovirus. The provirus may be transmitted from one cell generation to the next without causing lysis or destruction of the host cell.

5           A "self-inactivating 3' LTR" is a 3' long terminal repeat (LTR) that contains a mutation, substitution or deletion that prevents the LTR sequences from driving expression of a downstream gene. A copy of the U3 region from the 3' LTR acts as a template for the generation of both LTR's in the integrated provirus. Thus, when the 3' LTR with an inactivating deletion or mutation  
10 integrates as the 5' LTR of the provirus, no transcription from the 5' LTR is possible. This eliminates competition between the viral enhancer/promoter and any internal enhancer/promoter. Self-inactivating 3' LTRs are described, for example, in Zufferey et al. *J. Virol.* 72:9873-9880 (1998), Miyoshi et al. *J. Virol.* 72:8150-8157 and Iwakuma et al. *Virology* 261:120-132 (1999).

15

In one aspect of the invention, a recombinant retrovirus is used to deliver a transgene of interest to a cell, preferably an oocyte or an embryonic cell, more preferably a one-cell embryo. The transgene, and any associated genetic elements, are thus integrated into the genome of the host cell as a  
20 provirus. The cell may then be allowed to develop into a transgenic animal.

The recombinant retrovirus used to deliver the transgene is preferably a modified lentivirus, and thus is able to infect both dividing and non-dividing cells. The recombinant retrovirus preferably comprises a modified lentiviral genome that includes the transgene. Further, the modified lentiviral genome  
25 preferably lacks endogenous genes for proteins required for viral replication, thus preventing replication in the transgenic animal. The required proteins are provided in *trans* in the packaging cell line during production of the recombinant retrovirus, as described below.

In the preferred embodiment the transgene is incorporated into a viral  
30 construct that comprises an intact retroviral 5' LTR and a self-inactivating 3' LTR. The viral construct is preferably introduced into a packaging cell line that packages viral genomic RNA based on the viral construct into viral

particles with the desired host specificity. Viral particles are collected and allowed to infect the host cell. Each of these aspects is described in detail below.

5     The Viral Construct

The viral construct is a nucleotide sequence that comprises sequences necessary for the production of recombinant retrovirus in a packaging cell. In one embodiment the viral construct additionally comprises genetic elements that allow for the desired expression of a gene of interest in the host.

10     Generation of the viral construct can be accomplished using any suitable genetic engineering techniques well known in the art, including, without limitation, the standard techniques of restriction endonuclease digestion, ligation, transformation, plasmid purification, and DNA sequencing, for example as described in Sambrook et al. (Molecular Cloning: A  
15     Laboratory Manual. Cold Spring Harbor Laboratory Press, N.Y. (1989)), Coffin et al. (Retroviruses. Cold Spring Harbor Laboratory Press, N.Y. (1997)) and "RNA Viruses: A Practical Approach" (Alan J. Cann, Ed., Oxford University Press, (2000)).

20     The viral construct may incorporate sequences from the genome of any known organism. The sequences may be incorporated in their native form or may be modified in any way. For example, the sequences may comprise insertions, deletions or substitutions. In the preferred embodiment the viral construct comprises sequences from a lentivirus genome, such as the HIV genome or the SIV genome.

25     The viral construct preferably comprises sequences from the 5' and 3' LTRs of a lentivirus. More preferably the viral construct comprises the R and U5 sequences from the 5' LTR of a lentivirus and an inactivated or self-inactivating 3' LTR from a lentivirus. The LTR sequences may be LTR sequences from any lentivirus from any species. For example, they may be  
30     LTR sequences from HIV, SIV, FIV or BIV. Preferably the LTR sequences are HIV LTR sequences.



The viral construct preferably comprises an inactivated or self-inactivating 3' LTR. The 3' LTR may be made self-inactivating by any method known in the art. In the preferred embodiment the U3 element of the 3' LTR contains a deletion of its enhancer sequence, preferably the TATA box, Sp1 and NF-kappa B sites. As a result of the self-inactivating 3' LTR, the provirus that is integrated into the host cell genome will comprise an inactivated 5' LTR.

Optionally, the U3 sequence from the lentiviral 5' LTR may be replaced with a promoter sequence in the viral construct. This may increase the titer of virus recovered from the packaging cell line. An enhancer sequence may also be included. Any enhancer/promoter combination that increases expression of the viral RNA genome in the packaging cell line may be used. In the preferred embodiment the CMV enhancer/promoter sequence is used.

In one embodiment the viral construct comprises a gene that encodes a protein that is desirably expressed in one or more cells of a transgenic animal. Preferably the gene of interest is located between the 5' LTR and 3' LTR sequences. Further, the gene of interest is preferably in a functional relationship with other genetic elements, for example transcription regulatory sequences such as promoters and/or enhancers, to regulate expression of the gene of interest in a particular manner once the transgene is incorporated into the host genome. In certain embodiments, the useful transcriptional regulatory sequences are those that are highly regulated with respect to activity, both temporally and spatially.

Preferably the gene of interest is in a functional relationship with internal promoter/enhancer regulatory sequences. An "internal" promoter/enhancer is one that is located between the 5' LTR and the 3' LTR sequences in the viral construct and is operably linked to the gene that is desirably expressed.

The internal promoter/enhancer may be any promoter, enhancer or promoter/enhancer combination known to increase expression of a gene with which it is in a functional relationship. A "functional relationship" and

“operably linked” mean, without limitation, that the gene is in the correct location and orientation with respect to the promoter and/or enhancer that expression of the gene will be affected when the promoter and/or enhancer is contacted with the appropriate molecules.

5           The internal promoter/enhancer is preferably selected based on the desired expression pattern of the gene of interest and the specific properties of known promoters/enhancers. Thus, the internal promoter may be a constitutive promoter. Non-limiting examples of constitutive promoters that may be used include the promoter for ubiquitin, CMV (Karasuyama et al J. Exp. Med. 169:13 (1989),  $\beta$ -actin (Gunning et al. Proc. Natl. Acad. Sci. USA 84:4831-4835 (1987) and pgk (see, for example, Adra et al. Gene 60:65-74 (1987), Singer-Sam et al. Gene 32:409-417 (1984) and Dobson et al. Nucleic Acids Res. 10:2635-2637 (1982)). Alternatively, the promoter may be a tissue specific promoter. Several non-limiting examples of tissue specific promoters that may be used include lck (see, for example, Garvin et al. Mol. Cell Biol. 8:3058-3064 (1988) and Takadera et al. Mol. Cell Biol. 9:2173-2180 (1989)), myogenin (Yee et al. Genes and Development 7:1277-1289 (1993), and thyl (Gundersen et al. Gene 113:207-214 (1992). In addition, promoters may be selected to allow for inducible expression of the transgene. A number of systems for inducible expression using such a promoter are known in the art, including the tetracycline responsive system and the *lac* operator-repressor system. It is also contemplated that a combination of promoters may be used to obtain the desired expression of the gene of interest. The skilled artisan will be able to select a promoter based on the desired expression pattern of the gene in the resulting transgenic animal.

25           An internal enhancer may also be present in the viral construct to increase expression of the gene of interest. For example the CMV enhancer (Karasuyama et al J. Exp. Med. 169:13 (1989) may be used in combination with the chicken  $\beta$ -actin promoter. Again, one of skill in the art will be able to select the appropriate enhancer based on the desired expression pattern.

30           The gene of interest is not limited in any way and includes any gene that the skilled practitioner desires to have integrated and/or expressed in a

transgenic animal. For example, the gene of interest may be one that encodes a protein that modifies a physical characteristic of the transgenic animal, such as a protein that modifies size, growth, or tissue composition. In another example the gene of interest may encode a protein of commercial value that may be harvested from the transgenic animal.

In addition, more than one gene of interest may be placed in functional relationship with the internal promoter. For example a gene encoding a marker protein may be placed after the primary gene of interest to allow for identification of cells that are expressing the desired protein. In one embodiment a fluorescent marker protein, preferably green fluorescent protein (GFP), is incorporated into the construct along with the gene of interest. If a second reporter gene is included, an internal ribosomal entry site (IRES) sequence is also preferably included. The IRES sequence may facilitate the expression of the reporter gene

The viral construct may also contain additional genetic elements. The types of elements that may be included in the construct are not limited in any way and will be chosen by the skilled practitioner to achieve a particular result. For example, a signal that facilitates nuclear entry of the viral genome in the target cell may be included. An example of such a signal is the HIV-1 flap signal.

Further, elements may be included that facilitate the characterization of the provirus integration site in the genome of the animal. For example, a tRNA amber suppressor sequence may be included in the construct.

In addition, the construct may contain one or more genetic elements designed to enhance expression of the gene of interest. For example, a woodchuck hepatitis virus responsive element (WRE) may be placed into the construct (Zufferey et al. *J. Virol.* 74:3668-3681 (1999); Deglon et al. *Hum. Gene Ther.* 11:179-190 (2000)).

A chicken  $\beta$ -globin insulator may also be included in the viral construct. This element has been shown to reduce the chance of silencing the integrated provirus in the transgenic animal due to methylation and heterochromatinization effects. In addition, the insulator may shield the

internal enhancer, promoter and exogenous gene from positive or negative positional effects from surrounding DNA at the integration site on the chromosome.

Any additional genetic elements are preferably inserted 3' of the gene of interest.

In a specific embodiment, the viral vector comprises: a cytomegalovirus (CMV) enhancer/promoter sequence; the R and U5 sequences from the HIV 5' LTR; the HIV-1 flap signal; an internal enhancer; an internal promoter; a gene of interest; the woodchuck hepatitis virus responsive element; a tRNA amber suppressor sequence; a U3 element with a deletion of its enhancer sequence; the chicken  $\beta$ -globin insulator; and the R and U5 sequences of the 3' HIV LTR.

The viral construct is preferably cloned into a plasmid that may be transfected into a packaging cell line. The preferred plasmid preferably comprises sequences useful for replication of the plasmid in bacteria.

#### Production of Virus

Any method known in the art may be used to produce infectious retroviral particles whose genome comprises an RNA copy of the viral construct described above.

Preferably, the viral construct is introduced into a packaging cell line. The packaging cell line provides the viral proteins that are required in *trans* for the packaging of the viral genomic RNA into viral particles. The packaging cell line may be any cell line that is capable of expressing retroviral proteins. Preferred packaging cell lines include 293 (ATCC CCL X), HeLa (ATCC CCL 2), D17 (ATCC CCL 183), MDCK (ATCC CCL 34), BHK (ATCC CCL-10) and Cf2Th (ATCC CRL 1430). The most preferable cell line is the 293 cell line.

The packaging cell line may stably express the necessary viral proteins. Such a packaging cell line is described, for example, in U.S. Patent No. 6,218,181. Alternatively a packaging cell line may be transiently transfected

with plasmids comprising nucleic acid that encodes the necessary viral proteins.

In one embodiment a packaging cell line that stably expresses the viral proteins required for packaging the RNA genome is transfected with a plasmid comprising the viral construct described above.

In another embodiment a packaging cell line that does not stably express the necessary viral proteins is co-transfected with two or more plasmids essentially as described in Yee et al. (*Methods Cell. Biol.* 43A, 99-112 (1994)). One of the plasmids comprises the viral construct comprising the transgene. The other plasmid(s) comprises nucleic acid encoding the proteins necessary to allow the cells to produce functional virus that is able to infect the desired host cell.

The packaging cell line may not express envelope gene products. In this case the packaging cell line will package the viral genome into particles that lack an envelope protein. As the envelope protein is responsible, in part, for the host range of the viral particles, the viruses are preferably pseudotyped. Thus the packaging cell line is preferably transfected with a plasmid comprising sequences encoding a membrane-associated protein that will permit entry of the virus into a host cell. One of skill in the art will be able to choose the appropriate pseudotype for the host cell that is to be used. For example, in one embodiment the viruses are pseudotyped with the vesicular stomatitis virus envelope glycoprotein (VSVg). In another embodiment, a mutant ecotropic envelope protein is used, such as the ecotropic envelope protein 4.17 (Powell et al. *Nature Biotechnology* 18(12):1279-1282 (2000)). In addition to conferring a specific host range the pseudotype may permit the virus to be concentrated to a very high titer and may enhance safety by preventing the virus from infecting undesired cell types.

In the preferred embodiment a packaging cell line that does not stably express viral proteins is transfected with the viral construct, a second vector comprising the HIV-1 packaging vector with the *env*, *nef*, 5'LTR, 3'LTR and *vpu* sequences deleted, and a third vector encoding an envelope glycoprotein. Preferably the third vector encodes the VSVg envelope glycoprotein.

The recombinant virus is then preferably purified from the packaging cells, titered and diluted to the desired concentration.

#### Transgenic Animals

5 In order to make transgenic animals, an oocyte or one or more embryonic cells are infected with the recombinant virus produced as described above. One of skill in the art will recognize that the method of infection and the treatment of the cell following infection will depend upon the type of animal from which the cell is obtained. For example, mammalian cells are  
10 preferably implanted in a pseudopregnant female following infection while for the generation of transgenic birds or fish, the virus is preferably delivered to a laid egg and thus implantation is not required.

While early methods of making transgenic animals required the cells to be rapidly dividing, there is no such requirement in the methods of the present invention. Thus the cell may be contacted at any point in development. In the  
15 preferred embodiment, a zygote is contacted with the recombinant virus.

The cells to be infected with the virus may be obtained by any method known in the art and appropriate for the specific species in which it is desired to make a transgenic animal. For example, the recovery of fertilized mouse  
20 oocytes is described in Hogan et al. (Manipulating the Mouse Embryo: A Laboratory Manual. 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, NY (1994)). A method for obtaining fertilized rat oocytes is described in Armstrong et al. (Biol. Reprod. 39,511-518 (1998)).

It is not necessary that the cells be contacted after fertilization. In one  
25 embodiment, the virus is delivered to unfertilized ova. Development may then be initialized, for example by *in vitro* fertilization.

#### Delivery of the Virus

The virus may be delivered to the cell in any way that allows the virus  
30 to infect the cell. Preferably the virus is allowed to contact the cell membrane. Two preferred methods of delivering the virus to mammalian cells, injection and direct contact, are described below.

### Injection

In a first embodiment the virus is injected into the perivitelline space between the zona pellucida and the cell membrane of a single-cell zygote. Preferably less than 50 picoliters of viral suspension is injected, more  
5 preferably less than 25 picoliters and even more preferably about 10 picoliters.

The virus is preferably present in a viral suspension and may be injected by any method known in the art. The viral suspension is preferably injected through a hydraulic injector. More preferably a glass micropipette is used to inject the virus. In one embodiment a micropipette is prepared by  
10 pulling borosilicate glass capillary on a pipette puller. The tip is preferably opened and beveled to approximately 10  $\mu\text{m}$ . The lentiviral suspension may be loaded into the micropipette from the tip using gentle negative pressure.

In one embodiment the cell is stabilized with a holding pipette mounted on a micromanipulator, such as by gentle negative pressure against a  
15 fire-polished pipette, and a second micromanipulator is used to direct the tip of a micropipette into the space between the zona pellucida and the cell membrane, where the virus is injected.

### Direct contact

In another embodiment the zona pellucida is removed from the cell to  
20 produce a denuded embryo and the cell membrane is contacted with the virus. The zona pellucida may be removed by any method known in the art. Preferably it is removed by enzymatic treatment. For example, treatment with pronase may be used to remove the zona pellucida while the cell membrane is kept intact. Alternatively, the cell may be placed in media at pH at which the  
25 zona pellucida dissolves while the cell membrane remains intact. For example the cell may be incubated in an acidic Tyrode's solution at room temperature for several minutes. Once the zona pellucida is removed, any method that allows for the virus to contact the cell membrane may be used. Preferably, the cell is incubated in a solution containing the virus. Even more preferably, the  
30 solution is media that facilitates survival of the cell.

In this embodiment, the cells are preferably contacted with the virus in culture plates. The virus may be suspended in media and added to the wells of

a multi-well culture plate. The cells may then be plated in the individual wells. The media containing the virus may be added prior to the plating of the cells or after the cells have been plated. Preferably individual cells are incubated in approximately 10  $\mu$ l of media. However, any amount of media  
5 may be used as long as an appropriate concentration of virus in the media is maintained such that infection of the host cell occurs.

The cells are preferably incubated with the virus for a sufficient amount of time to allow the virus to infect the cells. Preferably the cells are incubated with virus for at least 1 hour, more preferably at least 5 hours and  
10 even more preferably at least 10 hours.

Both the injection and direct contact embodiments may advantageously be scaled up to allow high throughput transgenesis. Because of the relative simplicity of the injection technique, it is possible to inject many embryos rapidly. For example, it is possible to inject more than 200 fertilized oocytes  
15 in less than one hour. With regard to the direct contact embodiment, any number of embryos may be incubated in the viral suspension simultaneously. This may be accomplished, for example, by plating the desired number of single-cell zygotes in multi-well tissue culture plates containing the virus suspended in media appropriate for the survival and growth of the cells.

20 In both embodiments, any concentration of virus that is sufficient to infect the cell may be used. Preferably the concentration is at least 1 pfu/ $\mu$ l, more preferably at least 10 pfu/ $\mu$ l, even more preferably at least 400 pfu/ $\mu$ l and even more preferably at least  $1 \times 10^4$  pfu/ $\mu$ l.

Following infection with the virus, the cells are preferably implanted in  
25 an animal. More preferably cells infected with the virus are implanted in pseudo-pregnant animals of the same species from which the infected cells were obtained. Methods of creating pseudo-pregnancy in animals and implanting embryos are well known in the art and are described, for example, in Hogan et al. (Manipulating the Mouse Embryo: A Laboratory Manual. 2<sup>nd</sup>  
30 ed. Cold Spring Harbor Laboratory Press, NY (1994)).

In the preferred embodiment early stage embryos (approximately 0 - 2.5 days p.c.) still with an intact zona pellucida are transferred to the oviduct



of timed pseudopregnant female (preferably 0.5 days p.c.), while embryos that have reached the blastocyst stage are transferred to the uterus of timed pseudopregnant females (preferably 2.5 days p.c.). Denuded embryos are preferably cultured in vitro until they reach the morula or blastocyst stage (48  
5 to 72 hours in culture), and are then implanted into appropriately timed pseudopregnant females.

The embryos and resulting animals may be analyzed, for example for integration of the transgene, the number of copies of the transgene that integrated, the location of the integration, the ability to transmit the transgene  
10 to progeny and expression of the transgene. Such analysis may be carried out at any time and may be carried out by any methods known in the art. Standard techniques are described, for example, in Hogan et al. (supra).

The methods of infecting cells disclosed above do not depend upon species-specific characteristics of the cells. As a result, they are readily  
15 extended to all mammalian species.

Initial experiments with mice indicate that of those animals that develop to full term, 80-90% carried at least one copy of the transgene and that, of these, approximately 85% express the gene of interest. Of the transgenic animals about 25% carry only 1 or 2 copies of the transgene. The  
20 highest number of proviral insertions observed was about 30. Of the animals that carried only 1 or 2 copies of the transgene, about 80% expressed the gene of interest.

As discussed above, the modified retrovirus can be pseudotyped to  
25 confer upon it a broad host range. One of skill in the art would also be aware of appropriate internal promoters to achieve the desired expression of a gene of interest in a particular animal species. Thus, one of skill in the art will be able to modify the method of infecting cells to create transgenic animals of any species.

30 For example, transgenic birds are created by delivering a modified retrovirus, as described above, to the primordial germ cells of early stage avian embryos. In one embodiment, freshly laid eggs are obtained and placed in a

temperature controlled, humidified incubator. Preferably, the embryonic blastodisc in the egg is gradually rotated to lie on top of the yolk. This may be accomplished by any method known in the art, such as by gently rocking the egg regularly, preferably every 15 minutes. Approximately 36 hours later, the modified retrovirus is delivered into the space between the embryonic disk and the perivitelline membrane. Preferably about 50 nL of viral solution is delivered, more preferably about 100 nL of viral solution is delivered, and even more preferably about 200 nL of viral solution is delivered. The viral solution may be delivered by any method known in the art for delivering compositions to the inside of an egg. In the preferred embodiment a window is opened in the shell, the viral solution is injected through the window and the shell window is closed. The eggs are preferably incubated until hatching. The eggs will hatch after approximately 20 days, depending upon the particular avian species from which they are obtained. Hatched chicks are preferably raised to sexual maturity and mated. The transgenic offspring of the founder animals may be identified by any method known in the art, such as Southern blot, PCR and expression analysis.

In another embodiment, transgenic fish are created by delivering the modified retrovirus, described above, to single-cell fish embryos. Fertilized fish eggs are collected by any method known in the art. The modified retrovirus is then preferably delivered to the space between the chorion and the cell membrane. This may be accomplished, for example, by loading the modified retrovirus in solution into a glass pipette. The pipette may then be used to pierce the chorion membrane and deliver the viral suspension. Preferably about 50 nL of viral solution is delivered, more preferably about 100 nL of viral solution is delivered, and even more preferably about 200 nL of viral solution is delivered. Injected embryos are preferably returned to a temperature-controlled water tank and allowed to mature. At sexual maturity the founder fish are preferably mated and their progeny analyzed for the presence of the transgene by any method known in the art.

As mentioned above, the methods of the present invention will also prove useful in techniques for identifying genes that are involved in specific biological processes, such as gene trap assays and large-scale mutagenesis screens.

5           Gene trap experiments allow the identification and cloning of a gene that is expressed in a particular tissue or cell type, and/or at a particular time, based solely on its pattern of expression. Gene trapping relies on the capture of the splicing donor of an mRNA by ectopically inserting a downstream splice acceptor, in this case, carried within an integrated provirus. Gene  
10           trapping has been successfully used in several model systems, including the fruit fly *Drosophila*, mammalian cells in culture, and mouse ES cells (which have the advantage of being able to be used to derive mice afterwards for further analysis). Gene trapping in cell culture has the advantage of being fast and inexpensive, but is limited by the inability of the cells to differentiate into  
15           specific cell types. Thus, gene trapping experiments in mammalian cell lines in culture usually yield only housekeeping genes expressed non-specifically in any mammalian cell, or cell-specific genes that are only expressed by the particular cell line *in vitro*. Because cell lines often show incomplete degrees of differentiation, the complement of tissue-specific genes expressed by these  
20           cells is limited. Furthermore, there are many tissues for which representative cell lines do not exist.

          The use of the above-described recombinant lentiviral vectors for the purposes of gene trapping is facilitated by the self-inactivating mutation in the U3 enhancer element of the 3' LTR. The lack of transcriptional activity from  
25           the integrated 5' LTR ensures that any transcription of a reporter element in the provirus is driven by upstream regulatory sequences to the insertion that have been "trapped," rather than from the viral promoter itself.

          Thus, one embodiment of the present invention concerns a method of identifying genes that are expressed in a particular tissue and/or at a particular  
30           time during the development of an organism. A self-inactivating viral construct is made that preferably comprises a splice acceptor sequence and a sequence encoding a reporter gene. Modified retroviral particles are made

using the viral construct as described above and used to infect embryonic cells. Tissues from the founder animal or its progeny are analyzed for the presence of the reporter to determine the temporal and/or spatial pattern of expression. Messenger RNA is collected from the tissues of the animals that  
5 express the reporter protein in the time and place of interest. The "trapped" gene may then be identified by any method known in the art. Preferably, oligonucleotides that are complementary to the reporter gene may then used in a reverse transcription reaction to produce a cDNA that contains the sequences of the trapped gene that flank the provirus. The cDNA may then be cloned  
10 into a plasmid from which the gene may be identified by nucleotide sequencing.

Gene trap experiments are well known in the art and the skilled artisan would be able to choose the reporter gene, splice acceptor sequence and any other genetic elements that would be useful to include in the viral construct  
15 based on the specific analysis that they have undertaken. In addition, by modifying the viral constructs, the technique can be used to trap promoter or enhancer sequences. For promoter trap experiments, the reporter gene lacks any transcriptional regulatory elements, and is only expressed when the virus integrates next to an active promoter. For the enhancer trap, the reporter gene  
20 is positioned downstream of a minimal promoter that lacks transcriptional activity, and is only expressed when the virus integrates next to an active enhancer.

Another important paradigm by which biologists study gene function is to disrupt the function of an endogenous gene and, from the mutant phenotype  
25 that results, deduce the normal role of that gene in the organism. One way of isolating genes that are important in a particular biological process under study is to perform large-scale mutagenesis to generate animals that are phenotypically mutant in that process and then to isolate the gene that is disrupted in the mutant animal and that is thus responsible for the mutant  
30 phenotype. In most such experiments, either radiation or chemicals have been used to induce deletions or nonsense mutations. However, the genes carrying mutations induced by radiation or chemicals are difficult to isolate because no

handle is available with which to clone the gene. Rather, these mutations must be identified by positional cloning, a slow and painstaking process in which the mutation of interest is systematically mapped relative to known genetic markers in the genome, with the goal of gradually narrowing down and  
5 pinpointing the locus of the mutated gene.

In contrast, a powerful technique that has been used successfully in the fruit fly *Drosophila melanogaster* is that of insertional mutagenesis, in which genes are disrupted when an exogenous piece of DNA is inserted within the coding sequence of the gene. The great advantage of insertional mutagenesis  
10 is that, because the sequence of the exogenous disrupting DNA is known, one can directly clone out that piece of DNA and the flanking sequence that corresponds to the gene of interest that has been disrupted. Thus, in contrast to traditional positional cloning strategies used in chemical mutagenesis which may take up to 3 years after the isolation of the mutant, identification and  
15 characterization of the mutated gene of interest in an insertional mutagenesis strategy is reduced to just a week or so. The main limitation to the application of insertional mutagenesis in organisms other than the fruit fly is the lack of a DNA element, such as the transposon used in *Drosophila*, that is able to stably integrate and mark its position in the genome at the germline or one-cell  
20 embryo stage.

The lentiviral vectors described above can be effective tools for large-scale mutagenesis to identify genes involved in specific biological processes. The modified lentiviruses of the present invention are easily delivered to the germline, and pseudotyping of the viruses with an envelope glycoprotein, such  
25 as VSVg, allows the concentration of the virus to extremely high titers. Thus in one embodiment mutagenesis is achieved by delivery of the modified retrovirus to the cell membrane of embryonic cells.

The ability of transgenic animals made by the methods of the present invention to express a gene of interest at high levels suggests that the  
30 integrated proviruses are not silenced by methylation. Previous mutagenesis screens using MoLV-based retroviruses have been limited by the observation that, in addition to the provirus, flanking genomic sequences are frequently

found to be methylated and inactivated. This methylation complicates the analysis because it becomes difficult to distinguish whether the mutant phenotype is due to the disruption of the gene into which the provirus has inserted, or due to the inactivation of any one of several surrounding genes by methylation.

By delivering modified lentiviruses to embryos according to the methods of the present invention, insertional mutagenesis strategies can be applied to any animal species, including model genetic organisms such as *Xenopus*, zebrafish, mouse, rat, and zebra finch. Early-stage embryos, consisting of several cells, will preferably be targeted because the resulting mosaicism increases the number of unique mutagenic events that can be screened.

The modified lentiviruses integrate randomly into the genome of the target zygotes, including that of the germ cells. Thus, some proportion will disrupt coding sequences. The embryos are preferably raised to sexual maturity, mated, and the progeny are screened for mutant phenotypes of interest. Once a mutant is identified, selective breeding using standard methods is preferably used to isolate the particular insertion(s) responsible for the phenotype.

Once a mutant line is established, the mutated locus is preferably identified using the provirus as a handle for cloning. In one embodiment, an origin of replication and antibiotic resistance gene is included in the viral construct. In this embodiment, genomic DNA from the mutant is preferably isolated, randomly cleaved with an appropriate restriction enzyme, and the linear fragments circularized by ligation. The ligation mixture is then transformed into bacterial cells and plated on antibiotic plates. The plasmid DNA from any colonies that arise is isolated and preferably used as a template for inverse PCR with oppositely oriented, adjacent primers complementary to sequences in the provirus. The amplified DNA molecule(s) is then sequenced to acquire the flanking regions to the integration site, corresponding to the gene(s) mutated.

In another embodiment, inclusion of the tRNA amber suppressor sequence in the provirus allows for the rapid generation of genomic libraries containing the flanking regions of the integration loci, representing the disrupted gene(s). Once these flanking regions are sequenced, they can be compared against the genomic sequence database for that animal to determine candidate gene(s) of interest.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

#### **Example 1**

Transgenic mice were generated that expressed a heterologous protein, green fluorescent protein (GFP). GFP expression was controlled by manipulating the genetic elements in the viral construct employed to create the transgenic mice. For example, a viral construct, FUGW, comprising a ubiquitous promoter was used to produce transgenic mice that expressed GFP in every cell. Inclusion of a nuclear localization signal produced transgenic mice that had GFP localized in the nucleus of their cells. A viral construct with a lymphocyte specific promoter produced mice that expressed GFP in lymphocytes, while a viral construct with a muscle specific promoter produced mice that specifically expressed GFP in muscle cells.

#### **A. Viral Constructs**

A viral construct according to the present invention was created using the HR'CS-G plasmid (Miyoshi H, Blomer U, Takahashi M, Gage FH, Verma IM. *J Virol.* 72(10):8150-7 (1998)). This plasmid is based on the HIV-1 HXB2 proviral DNA (see U.S. Patent No. 6,013,516).

1. **Generation of a vector expressing GFP from a ubiquitous promoter, FUGW**

5 The HIV-1 flap sequence (SEQ ID NO: 1) was inserted into the HR'CS-G vector. A 147 base pair sequence containing the flap region (Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., Charneau, P. *Cell*. 101(2),173-185 (2000)) was PCR amplified from a plasmid encoding the genome of HIV NL4.3. The 5' PCR primer encoded BglII and PacI sites. The 3' PCR primer contained the BamHI site. The resulting PCR  
10 product was digested with BglII and BamHI enzymes and inserted into the BamHI site of the HR'CS-G vector. The resulting plasmid was called Hflap.

The ubiquitin promoter (SEQ ID NO: 2) was then inserted into the Hflap plasmid. The 1.2 Kb sequence encoding the human polyubiquitin C promoter was excised with BglII and BamHI enzymes and inserted into the  
15 BamHI site of Hflap. The resulting plasmid was called HflapUbi.

A multi-cloning site was then inserted into HflapUbi. Two oligonucleotides were designed that encoded the following restriction sites: BamHI HpaI XhoI AscI EcoRI BglII. The oligos were hybridized and inserted into the BamHI site of HflapUbi. The resulting plasmid was called FUMCS.

20 A nucleic acid sequence encoding GFP (SEQ ID NO: 3) was then inserted into HflapUbi. The 700 base pair sequence of GFP was digested with BamHI and XhoI and inserted into the XhoI site of HflapUbi, generating HflapUbiG. The resulting plasmid was called FUG.

The woodchuck hepatitis virus regulator element (WRE; SEQ ID NO: 4) was then inserted into HflapUbiG. The 500 bp sequence of WRE (Zufferey, R., Donello, J.E., Trono, D., Hope T.J. (1999). *J. Virol.* 73(4), 2886-92) was excised with SalI and XhoI and inserted into the XhoI site of HflapUbiG, generating HflapUbiGWRE (SEQ ID NO: 8). The resulting  
25 plasmid was called FUGW. A map of the FUGW viral vector is presented in  
30 Figure 1A.



2. **Generation of a vector expressing nuclear-localized GFP from a ubiquitous promoter, FUH2BGW**

In order to get specific nuclear localization of GFP, the histone 2B-GFP fusion sequence H2BGFP was cloned into FUGW. The histone 2B-GFP sequence (SEQ ID NO: 5) was digested with SalI and NotI. Both sites were blunted with T4 DNA polymerase and inserted into the HpaI site of FUMCS. The resulting plasmid was called FUH2BGW.

3. **Generation of a vector expressing GFP from a lymphocyte-specific promoter, FlckGW**

To achieve lymphocyte specific GFP expression the murine lck promoter (SEQ ID NO: 6) was cloned into the FUGW vector. The ubiquitin promoter from FUGW was removed by excising with PacI and BamHI. The PacI site was blunted using T4 DNA polymerase and the lck promoter was excised with SpeI and BamHI. The SpeI site was blunted using T4 DNA polymerase. The lck promoter was inserted into the PacI and BamHI sites of FUGW. The resulting plasmid was called FlckGW.

4. **Generation of a vector expressing GFP from a muscle-specific promoter, FmyoH2BGW**

To achieve specific expression of GFP in the muscle of transgenic mice the myogenin promoter (SEQ ID NO: 7) was cloned into the FUGW construct. The mouse myogenin promoter was PCR amplified from a mouse genomic BAC. The 5' PCR primer encoded a PacI site, and the 3' PCR primer contained an XbaI site. The PCR product was digested with PacI and XbaI. The ubiquitin promoter was removed from the FUH2BGW vector by cutting with PacI and XbaI. The PacI-BamHI digested PCR product encoding the myogenin promoter was inserted into the PacI and XbaI sites of the FUH2BGW vector.

The constructs described above were then used to prepare recombinant lentivirus. Briefly, replication-incompetent viral vectors, based on the human immunodeficiency virus-1 (HIV-1), were pseudotyped with the vesicular stomatitis virus envelope glycoprotein (VSVg), permitting the virus to be

concentrated to very high titers and conferring upon the virus a broad host range. Pseudotyped lentiviruses were produced essentially as described in detail in Yee, J.K., Friedmann, T. & Burns, J.C. (1994). *Methods Cell Biol.* 43, 99-112; Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J.K. (1993). *Proc. Natl. Acad. Sci. USA.* 90, 8033-8037; and Yee, J.K., Miyanohara, A., LaPorte, P., Bouic, K., Burns, J.C., and Friedmann, T. (1994). *Proc. Natl. Acad. Sci. USA.* 91, 9564-9568. Briefly, human fibroblasts 293 cells were transfected with calcium phosphate/DNA coprecipitates of the following plasmids, as described in Gorman, C., Padmanabhan, R. and Howard, B.H. (1983). *Science.* 221, 551-553:

The viral transfer vector described above with self-inactivating LTR; the HIV-1 packaging vector  $\Delta 8.9$  (Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., and Trono, D. (1997). *Nat. Biotechnol.* 15(9), 871-875; Naldini, L., Blomer, U., Galloway, P., Ory, D., Mulligan, R., Gage, F.H., Verman, I.M., and Trono, D. (1996). *Science.* 272(5259), 263-67) with *env*, *vpr*, *vpu*, *vif*, *nef*, 5'LTR, 3'LTR, and  $\psi$  sequences deleted; and the VSVg envelope glycoprotein.

Viral supernatant was harvested 60 hours post-transfection, subjected to low-speed centrifugation to remove cell debris, filtered through a 0.45  $\mu$ m nitrocellulose membrane, spun at 25,000 rpm for 1.5 hours to concentrate, and resuspended in a small volume (one hundredth to one thousandth of the original volume) of phosphate-buffered saline (PBS), pH 7.4. The titer of the viral concentrate was approximately  $1 \times 10^6$  pfu/ $\mu$ l as determined in 293 human fibroblasts measured by the number of GFP-positive cells. The viral suspension was stored frozen at -80°C.

## **B. Production of transgenic mice and rats**

The lentivirus was used to produce transgenic mice and rats.

### **1. Superovulation and embryo collection**

Female mice and rats were superovulated with a combination of pregnant mare's serum (PMS) and human chorionic gonadotropin (hCG) as

described in Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. "Superovulation" refers to administering gonadotropins to female mammals prior to mating to increase the number of eggs that are ovulated. Prepubescent female mice (approximately 25 days of age and weighing between 12.5 and 14 grams) were injected intraperitoneally with 5 IU of PMS (Sigma G 4527, 25 IU/ml in 0.9% NaCl) between 1 and 3 p.m. on day -2, followed by 5 IU of HCG (Sigma C 8554, 25 IU/ml in 0.9% NaCl) 48 hours later on day 0. Prepubescent female rats between 28-30 days of age and weighing between 70 and 80 grams were injected intraperitoneally with 25 IU of PMS between 1 and 3 p.m. on day -2, followed by 5 IU of HCG 48 hours later on day 0. For both rats and mice, hormone-treated females were caged overnight with fertile males (2-3 months of age) to mate. On the morning of day 1, females were checked for copulation plugs.

Female mice were sacrificed for embryo collection around 10 a.m. on the morning of day 1, while female rats were sacrificed for collection around 1 p.m. on the afternoon of day 1. Embryos were collected from mice and rats essentially following the procedure described in Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994) *Manipulating the Mouse Embryo: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. Briefly, animals were sacrificed by CO<sub>2</sub> inhalation, and the oviducts were excised and transferred to a dish containing M2 medium at room temperature. Newly fertilized embryos, enclosed by cumulus mass cells, were released from the swollen ampullae (the upper portion of the oviduct) by gently tugging and opening the walls of the ampullae with fine forceps. The embryos were then transferred to a dish containing a hyaluronidase solution (Sigma H 3884, 300 µg/ml in M2 medium), which enzymatically digested the cumulus cells, thus releasing the embryos. When the cumulus cells were shed, the embryos were transferred to fresh M2 medium to wash off the hyaluronidase solution and preserve the viability of the embryos. In rats, the cumulus cells were found to adhere tenaciously to the surface of some embryos and were difficult to

remove completely. Thus in some cases the subsequent experimental manipulations with the zygotes were carried out with some of the cumulus cells still adhering. This did not seem to affect the outcome. The embryos were then transferred to microdrops of M16 medium under mineral oil and cultured in a humidified 37°C incubator under 5% CO<sub>2</sub> until needed.

## 2. Delivery of lentiviruses to single-cell embryos

Lentiviruses were delivered to the fertilized oocytes on the same day of collection, targeting only single-cell zygotes to minimize mosaicism. Infection with lentivirus derived from the FUGW viral construct will lead to integration of the provirus locus diagrammed in Figure 1B. Two different methods were used to deliver the lentiviruses to the embryos:

### a. Microinjection of lentiviruses into the perivitelline space of single-cell embryos

Micropipettes were prepared by pulling borosilicate glass capillaries (1 mm O.D., 0.7 mm I.D.) on a Sutter Instruments pipette puller. The tip was cut at an angle to approximately 10 µm with a razor blade. The micropipette was then inserted into the pipette holder of a CellTram hydraulic injector (Eppendorf). The lentiviral concentrate prepared above was pipetted up and down to release any large aggregates of cellular debris. The virus was centrifuged at low speed in a tabletop microfuge (1000 rpm for 1 min.), and removed from the top. The viral suspension was then loaded into the micropipette from the tip using gentle negative pressure from the CellTram.

One-cell embryos were transferred to a microdrop of M2 medium on a slide and covered with mineral oil to maintain the osmolarity. The slide was mounted on the stage of an inverted light microscope, and the injection procedure was monitored under 400X magnification. Embryos were held in place against a fire-polished pipette using gentle negative pressure. The pipette holder with the virus was loaded onto a micromanipulator (Leitz). Using the micromanipulator to guide the pipette, the tip was pushed through the zona pellucida into the region between the zona pellucida and the oocyte cell membrane. Using gentle positive pressure, approximately 10 nanoliters of

the viral concentrate was delivered into the perivitelline space. The micropipette was then withdrawn from the zygote. After the injection, the embryos were sorted and those that were lysed, abnormal, or at the 2-cell stage were discarded. The remaining embryos were transferred to M16 microdrops under oil and cultured in a 37°C incubator under 5% CO<sub>2</sub> until implantation.

**b. Co-incubation of denuded single-cell embryos with lentiviruses**

The zona pellucida of the fertilized oocytes was removed by incubation in either an acidic Tyrode's solution (Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*. Cold Spring Harbor Laboratory Press) or a 0.5% pronase solution in M2 medium at 37°C in a humidified 5% CO<sub>2</sub> incubator for several minutes. When the zonae appeared to be dissolved, embryos were washed in excess M2 medium and then transferred into 10 µl microdrops of viral suspension under mineral oil. Embryos were cultured individually in separate microdrops to prevent them from adhering to one another. The viral suspension was diluted to various concentrations to roughly control the average number of proviral integrations expected per transgenic genome. Virus in the microdrops was diluted to  $2 \times 10^4$  pfu/µl, 400 pfu/µl, and 8 pfu/µl. Zygotes were incubated in the viral suspension for at least 4-6 hours before implantation to allow viral entry into the cell.

**3. Transfer of embryos into recipient females**

Timed pseudopregnant females to host the treated embryos were prepared by mating sexually mature females in estrus to vasectomized, mature males the night before the intended day of implantation. Appropriate females were selected from a colony of 30-40 females by taking vaginal smears and examining them for the cell types typical of the estrus phase. Males were vasectomized by tying off the vas deferens at two separate locations, approximately 5-6 mm apart, then cauterizing the intervening segment to sever the tube. Males were vasectomized at least 2 weeks prior to the mating to

ensure that all remaining sperm in the genital tract were dead at the time of mating.

Embryos infected with lentivirus were transferred into host females as soon as possible to achieve maximum rates of implantation. Early-stage embryos (0-2.5 days p.c.) with an intact zona pellucida were transferred to the oviduct of timed pseudopregnant females (0.5 days p.c.), while embryos that had reached the morula or blastocyst stage were transferred to the uterus of timed pseudopregnant females (2.5 days p.c.). In general, no more than 30 embryos were transferred bilaterally into the uterus. These procedures were carried out essentially as described in (Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*. Cold Spring Harbor Laboratory Press). Pregnancy and delivery of the transgenic litter was as usual.

15

### C. Analysis of transgenic animals

Animals in the resulting litters were analyzed for the presence of the transgene and the number of insertions of the transgene by standard Southern blot analysis (Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor laboratory Press.), cutting with PstI or BamHI and hybridizing against a GFP+WRE probe. For constitutive promoters, expression of GFP was determined by directly viewing the skin of the animals under a conventional epifluorescence microscope. Some transgenic animals that were scored as negative for expression were actually expressing the transgene at levels below that of detection by visual inspection with a fluorescent microscope. In such cases, western blot analysis revealed that animals in which GFP fluorescence was not detected by visual inspection did express the GFP protein in some tissues. Similarly, immunocytochemistry proved to be a more sensitive assay for determining expression. For the tissue-specific promoters, some proportion of the transgenic litter was sacrificed during development at embryonic stages, and the translucent embryos were checked for spatially regulated GFP expression under a fluorescent microscope. Expression results were confirmed by

30

histology. To test the ability of the founder animals to transmit the transgene to their progeny, animals positive by Southern analysis were outcrossed to wild-type animals, and their progeny scored for transgenesis and expression as described above.

5

#### **D. Results**

In one set of experiments one-cell mouse and rat embryos were injected in the perivitelline space with recombinant lentivirus as described above. In the first experiment, 17 founder mice developed to term from 78  
10 implanted embryos. Of these, 11 of the 17 founders expressed the transgene as determined by directly viewing the animals under an epifluorescence microscope. Further, 11 of 15 (two mice died prior to analysis), or approximately 73%, were found to carry the transgene by Southern blot analysis. The average number of insertions in the transgenic mice was 6.1.  
15 Several of the animals carried as few as 2 insertions. These results are presented in Figure 2.

In a second experiment, 56 founder mice developed to term from 119 implanted embryos. Of these 45, or about 80%, were found to express the transgene. Thus, in the two experiments 58 out of 73 founder mice,  
20 approximately 79.5%, carried the transgene. Figure 3 shows the Southern blot analysis of proviral transgene insertions in these founder mice.

All GFP-positive mice carried an integrated provirus, and all animals with two or more copies of the provirus expressed the transgene at levels detectable by direct viewing of GFP fluorescence. The intensity of GFP  
25 fluorescence correlated positively with copy number, as estimated qualitatively. All major tissues and organs, including skin, bone, skeletal muscle, cardiac muscle, lung, liver, thymus, spleen, stomach, intestine, kidney, brain, retina and gonads were GFP positive.

In a third experiment, five rats developed to term from embryos  
30 injected with lentivirus created from the FUGW construct. Two of the five rats were found to express the transgene as determined by brightfield and fluorescent images of the paws of the newborn rats (Figure 4A). Pup R4

expresses GFP in the paw, as well as in all other tissues and organs that were examined (Fig. 4A). Figure 4B shows the Southern blot analysis of the proviral insertions in these founder rats and indicates that pup R4 carries 4 copies of the proviral insert.

5           In a continuation of this experiment, out of 22 founder rat pups born from 130 implanted embryos, 13 (59.1%) carried one or more proviral insertions as determined by Southern blot analysis and 9 (40.9%) expressed GFP at levels detectable by directly viewing the skin under a fluorescent microscope. GFP positive founders were crossed to wild-type animals, and F1  
10 progeny rats carrying as few as one copy of the provirus expressed GFP, as determined by direct viewing with a fluorescent microscope, indicating that the GFP-expressing transgene is not silenced by transmission through the germline.

          In another set of experiments denuded mouse embryos were incubated  
15 in decreasing concentrations of recombinant lentivirus. A rough correlation was seen between the titer of virus in which embryos were incubated and the number of proviral insertions. At a 1:50 dilution from a stock of  $1 \times 10^6$  pfu/ $\mu$ l 5 founder mice that reached term (from 29 implanted embryos) were found to be transgenic. All of these animals carried at least 6 proviral  
20 insertions. The average number of insertions was 7.2. Of these, 4, or 80%, were found to express the transgene. At a 1:250 dilution five out of 7 founder mice that reached term (from 18 implanted embryos) were found to be transgenic and express the transgene. In these mice the average number of  
25 insertions was 3.8, with two of the animals carrying only one or two copies of the transgene. Finally, at a dilution of 1:1250 only one of the 8 founders (from 40 implanted embryos) was found to be transgenic, with a single insertion. This founder also expressed the transgene. Figure 5 shows the Southern blot analysis of proviral transgene insertions in these founder mice. Figure 6 shows GFP expression in one of the founder mice. A second trial with a  
30 1:250 dilution gave comparable results. Eight of 11 founder mice (from 59 implanted embryos) were transgenic, with seven expressing the transgene. The transgenic mice had an average of 2.6 insertions.



Following outcrossing to wild-type animals, progeny were analyzed for proviral transgene insertions by Southern blot (Figure 7) and for GFP expression by viewing under an epifluorescence microscope (Figure 8). As can be seen in Table 1, founder mice were able to transmit the transgene to their progeny. In Table 1, "PV" represents founder transgenics generated by injection of the lentivirus into the perivitelline space of one-cell embryos while "Co-inc" represents founder transgenics generated by co-incubation of the denuded embryos with lentivirus.

**Table 1**

Founder	No. insertions in founder	No. progeny	No. expressing
PV.13	6	7	6/7
PV.2	2	4	2/4
PV.10	2	7	1/7
PV.17	10	10	10/10
Co-inc.18	12	12	9/12
Co-inc.2	0	9	0/9

Ubiquitous GFP expression similar to that of the founder animals was observed in transgenic F1 progeny, indicating that the provirus was not inactivated through one round of gametogenesis and development. All animals carrying two or more insertions of the FUGW provirus expressed GFP at levels detectable by direct fluorescence. However, among transgenic lines carrying one proviral insertion, approximately half expressed the transgene at levels detectable by direct fluorescence. In one single insertion line in which GFP expression was not observed by direct viewing, GFP was detectable by Western blot analysis in some tissues (brain, testes) but not in others (heart, lung, liver, kidney, spleen).

In a further experiment, single-cell mouse zygotes were injected in the perivitelline space with recombinant lentivirus derived from the FMH2BGW viral construct described above. This construct comprises a histone2B-GFP fusion gene under the control of the myogenin promoter. The histone2B-GFP

reporter was used to concentrate the fluorescence in the nuclei, making the signal more intense. Zygotes were implanted in pseudopregnant female mice and then recovered from the uterus at embryonic day 11.5. As can be seen in Figure 9, all six founder-mice were transgenic as determined by Southern blot analysis of proviral transgene insertions. Of these, two animals were positive for tissue-specific GFP expression at embryonic day 11.5 (Figure 9).

Figure 10 shows that at embryonic day 11.5, GFP expression is localized to the somites and can be seen in the emerging muscles in the limb buds, eye and jaw. This expression pattern is consistent with myogenin expression at this stage of development. Figure 11 shows the results of immunofluorescence studies of sections of an embryonic day 11.5 embryo carrying the myogenin promoter driving GFP. Specific staining of somite tissues can be seen (Fig. 11). Figures 12 and 13 show further immunofluorescence studies of cross-sections of an E11.5-embryo carrying a myogenin promoter driving GFP.

Fifteen-day old animals, derived from FMH2BGW-infected zygotes showed GFP fluorescence in the nuclei of skeletal muscle in the tongue, limbs, chest and jaw, but not in cardiac or smooth muscle or other non-muscle tissues examined, reflecting the known specificity of myogenin expression. F1 progeny from three independent founders expressed histone2B-GFP exclusively in the skeletal muscle lineage. Furthermore, progeny carrying as few as one FMH2BGW proviral insertion expressed histone2B-GFP in the appropriate tissue types at high levels detectable by direct viewing with a fluorescent microscope.

In a further experiment, a viral vector containing GFP driven by the T-lymphocyte promoter *lck*, *Flck*GW was delivered to the perivitelline space by injection as described above. The resulting transgenic mice expressed GFP exclusively in the thymus.

**Example 2**

Transgenic birds, such as chicken or quail, may be made by the methods of the present invention.

5           Freshly laid chicken eggs (day 0) are placed in a temperature-controlled, humidified incubator at 38°C. The embryonic blastodisc is gradually rotated to lie on top of the yolk by gently rocking the eggs in the incubator every 15 minutes. A window is opened in the shell and the blastodisc is visualized in freshly laid eggs (0 hours post-laying) or stage X  
10           embryos (36 hours post-laying). VSV-pseudotyped lentiviral particles in solution are loaded into a glass capillary micropipette. To maximize the chances of targeting primordial germ cells, virus is injected in the anterior regions of the 0 hour embryos and in the gonadal anlage of the 36 hours embryos. Approximately 200 nL of viral solution are delivered into the space  
15           between the perivitelline membrane and the embryonic disk with the aid of a hydraulic injector. The shell window is then closed with a porous tape to allow gas exchange between the embryo and the incubator atmosphere. The embryos are then incubated without rocking. The eggs will hatch after approximately 20 days of incubation time. Hatched chicks are raised to sexual  
20           maturity and then mated. The eggs laid by the mated females are raised to hatching and the resulting transgenic chicks are identified, such as by Southern blot, PCR or expression analysis.

**Example 3**

25           Transgenic zebra finch were made by the methods of the present invention.

            Freshly laid zebra finch eggs (day 0) were placed in a temperature-controlled, humidified incubator at 38°C. The embryonic blastodisc was  
30           gradually rotated to lie on top of the yolk by gently rocking the eggs in the incubator every 15 minutes. A window was opened in the shell and the blastodisc was visualized. VSV-pseudotyped lentiviral particles in solution

were loaded into a glass capillary micropipette. The lentivirus was derived from the FUH2BGW viral construct described above. To maximize the chances of targeting primordial germ cells, virus was injected in the anterior regions of 0 hour embryos and in the gonadal anlage of 36 hour embryos. Approximately 200 nL of viral solution are delivered into the space between the perivitelline membrane and the embryonic disk with the aid of a hydraulic injector.

The shell window was closed with a porous tape to allow gas exchange between the embryo and the incubator atmosphere. The embryos were then incubated without rocking. Figure 15 shows H2B-GFP expression in the extraembryonic tissue. Figure 16 shows H2B-GFP expression inside of the zebra finch embryo, indicating that primordial germ cells carried and expressed the transgene.

The eggs will hatch after approximately 20 days of incubation time. Hatched chimeric chicks are raised to sexual maturity and then mated. The eggs laid by the mated females are raised to hatching and the resulting transgenic chicks are identified, such as by Southern blot, PCR or expression analysis.

#### **Example 4**

Transgenic fish may be made by the methods of the present invention. Breeding pairs of fish are placed in a water tank with a grooved bottom, where fertilized eggs are deposited. Fertilized eggs (zygotes) are collected and held in embryo medium on ice. Zygotes are aligned in grooves formed in a slab of agarose. A modified lentivirus, as described above, is loaded into a glass capillary micropipette. The chorion membrane surrounding the zygote is pierced with the glass micropipette and 200 nL of viral solution are delivered into the space between the zygotic membrane and the chorion. Injected zygotes are returned to a temperature-controlled water tank and allowed to mature. At sexual maturity, the founder fish are mated and their progeny

analyzed for the presence of the transgene, such as by Southern blot, PCR and protein analysis.

### **Example 5**

5

The modified lentivirus described above may also be used in gene trap experiments, such as in zebrafish. As discussed above, this technique allows the identification and cloning of a gene that is expressed in a particular tissue or cell type and/or at a particular time based solely on its pattern of expression.

10 Zebrafish is an ideal system for gene trapping for several reasons. First, embryonic development occurs externally, allowing for easy manipulation and viewing of the embryos. Furthermore, early stage zebrafish embryos are translucent, and the pigmentation can be further suppressed for several more days by incubating the embryo in a 0.003% solution of 1-phenyl-2-thiourea (PTU). The translucent property of zebrafish embryos facilitates the viewing

15 of a live fluorescent reporter to identify trapped genes expressed in spatial or temporal patterns of interest.

Self-inactivating lentiviral vectors are engineered to contain a gene trap element consisting of the following sequences: splice acceptor-IRES-GFP-poly A addition signal. This cassette is called SAIGP. The SAIGP element is inserted in a 3' to 5' orientation with respect to the viral LTR sequences, to prevent inappropriate splicing or termination of the viral genome during packaging. Zebrafish zygotes are injected with VSVg-pseudotyped, concentrated SAIGP lentivirus as described above. Fish are

20 raised to sexual maturity and mated. The progeny are viewed with a fluorescent microscope, and GFP-expressing individuals are separated for further analysis. GFP-positive animals are then analyzed with a confocal fluorescent microscope to determine the spatial and temporal pattern of expression. Messenger RNA is extracted from those tissues of the animal that

25 express GFP in the time and place of interest, and reverse transcription with oligonucleotides complementary to GFP yields a cDNA that should contain the sequences of the trapped gene that flank the provirus. The recovered

30

cDNA is subcloned into an appropriate bacterial plasmid, and the gene that has been trapped by the SAIGFP provirus is identified by sequencing the upstream regions of the cDNA.

5     **Example 6**

Virus particles generated from the FUGW vector were generated as described above. The virus particles were injected using the perivitelline injection method, also described above, into 4 fertilized rhesus monkey (*Macaca mulatta*) single cell embryos. Monkey oocytes can be fertilized  
10 directly with sperm or can be fertilized using the intracytoplasmic sperm injection (ICSI) method. Of the four embryos injected, 2 developed into blastocysts. Both blastocysts were green, evidencing expression of GFP. In the injected blastocysts, cells in the trophectoder (TE) and the inner cell mass (ICM) both were green. Non-injected control embryos were not green.  
15 Transformed embryos are transferred to host mothers for gestation. After approximately 150 to 175 days, a newborn rhesus monkey is delivered which expresses GFP throughout. Confirmation of the presence of the transgene and expression in various tissues is carried out as described above.

WHAT IS CLAIMED IS:

1. A method of producing a transgenic animal comprising:  
transfecting a packaging cell line with a retroviral construct;  
recovering recombinant retrovirus from the packaging cell line;  
5 and  
infecting an embryonic cell with the recombinant retrovirus,  
wherein the retroviral construct comprises the R and U5 sequences  
from a 5' lentiviral long terminal repeat (LTR) and a self-inactivating  
lentiviral 3' LTR.
- 10 2. The method of claim 1 wherein said retroviral construct  
additionally comprises a gene of interest.
3. The method of claim 2 wherein said retroviral construct  
additionally comprises an internal promoter.
4. The method of claim 3 wherein said transgenic animal expresses  
15 the gene of interest.
5. The method of claim 1 wherein said packaging cell line is a 293  
cell line.
6. The method of claim 1 wherein the 5' LTR sequences are from  
HIV.
- 20 7. The method of claim 1 wherein the self-inactivating 3' LTR  
comprises a U3 element with a deletion of its enhancer sequence.
8. The method of claim 7 wherein the self-inactivating 3' LTR is a  
modified HIV 3' LTR.
9. The method of claim 1 wherein the recombinant retrovirus is  
25 pseudotyped.
10. The method of claim 9 wherein the recombinant retrovirus is  
pseudotyped with the vesicular stomatitis virus envelope glycoprotein.
11. The method of claim 9 wherein the recombinant retrovirus is  
pseudotyped with a mutant ecotropic envelope protein.
- 30 12. The method of claim 1 wherein the viral construct additionally  
comprises a promoter operably linked to the R and U5 5' LTR sequences.
13. The method of claim 12 wherein the promoter is a CMV promoter.

14. The method of claim 12 wherein the viral construct additionally comprises an enhancer operably linked to the promoter.

15. The method of claim 14 wherein the enhancer and promoter are CMV sequences.

5           16. The method of claim 1 wherein the viral construct additionally comprises the woodchuck hepatitis virus enhancer element sequence.

17. The method of claim 1 wherein the viral construct additionally comprises a tRNA amber suppressor sequence.

10           18. The method of claim 3 wherein the viral construct additionally comprises a reporter gene operably linked to the internal promoter.

19. The method of claim 18 wherein the reporter gene encodes a fluorescent protein.

20. The method of claim 19 wherein said fluorescent protein is green fluorescent protein.

15           21. The method of claim 3 wherein the internal promoter is a ubiquitous promoter.

22. The method of claim 21 wherein said ubiquitous promoter is selected from the group consisting of the ubiquitin promoter, the CMV  $\beta$ -actin promoter and the pgk promoter.

20           23. The method of claim 3 wherein the internal promoter is a tissue specific promoter.

24. The method of claim 23 wherein said tissue specific promoter is selected from the group consisting of the lck promoter, the myogenin promoter and the thyl promoter.

25           25. The method of claim 1 wherein infecting an embryonic cell comprises injecting the recombinant retrovirus between the zona pellucida and the cell membrane of a mammalian embryonic cell.

26. The method of claim 1 wherein infecting an embryonic cell comprises removing the zona pellucida from a mammalian embryonic cell and incubating the cell in solution containing the recombinant retrovirus.

30           27. The method of claim 26 wherein the zona pellucida is removed by enzymatic digestion.



28. A method of producing a transgenic mammal comprising the following steps:

- a) removing the zona pellucida from an embryonic cell;
- b) contacting the embryonic cell with a modified retrovirus; and
- c) implanting the embryonic cell in a pseudo-pregnant female.

29. The method of claim 28 wherein the embryonic cell is a one-cell embryo.

30. The method of claim 28 wherein the modified retrovirus is a modified lentivirus.

31. The method of claim 30 wherein the modified lentivirus is produced by transfecting a packaging cell line with a viral construct.

32. The method of claim 31 wherein the viral construct comprises the R and U5 sequences from a lentiviral 5' LTR, an internal promoter, a gene of interest, and a self-inactivating lentiviral 3' LTR.

33. The method of claim 28 wherein the embryonic cell is contacted with the modified retrovirus for at least 5 hours.

34. A method of producing a transgenic animal comprising injecting a modified lentivirus into the perivitelline space of an embryonic cell.

35. A transgenic animal produced by any one of claim 1, 28 or 34.

36. A transgenic animal whose genome comprises proviral DNA, said proviral DNA comprising a self-inactivating 3' lentiviral LTR.

37. The transgenic animal of claim 36 wherein the self-inactivating 3' LTR is a self-inactivating 3' HIV LTR.

38. The transgenic animal of claim 36 wherein the self-inactivating 3' HIV LTR comprises a U3 element with a deletion of its enhancer sequence.

39. A viral construct comprising the CMV enhancer/promoter sequence fused to the R and U5 sequences from the 5' HIV LTR, the HIV-1 flap signal, an internal enhancer/promoter, an exogenous gene of interest, the woodchuck hepatitis virus responsive element, a tRNA amber suppressor

sequence, a U3 element with a deletion of its enhancer sequence, the chicken  $\beta$ -globin insulator and the R and U5 sequences of the 3' HIV LTR.

40. A packaging cell comprising the viral construct of claim 39.

41. The packaging cell of claim 40 additionally comprising a second  
5 vector comprising the HIV-1 packaging vector with the *env*, *nef*, 5'LTR, 3'LTR and *vpu* sequences deleted.

42. The packaging cell of claim 41 additionally comprising a third vector encoding an envelope glycoprotein.

43. The packaging cell of claim 42 wherein the third vector encodes  
10 the VSVg envelope glycoprotein.

44. A method of isolating a gene that is expressed in a particular tissue comprising:

15 infecting embryonic cells with a modified retrovirus whose genome comprises a self-inactivating lentiviral 3' LTR, a splice acceptor sequence and a reporter gene;

isolating tissues from the embryo or mature animal that express the reporter gene; and

identifying the gene sequences that flank the provirus.

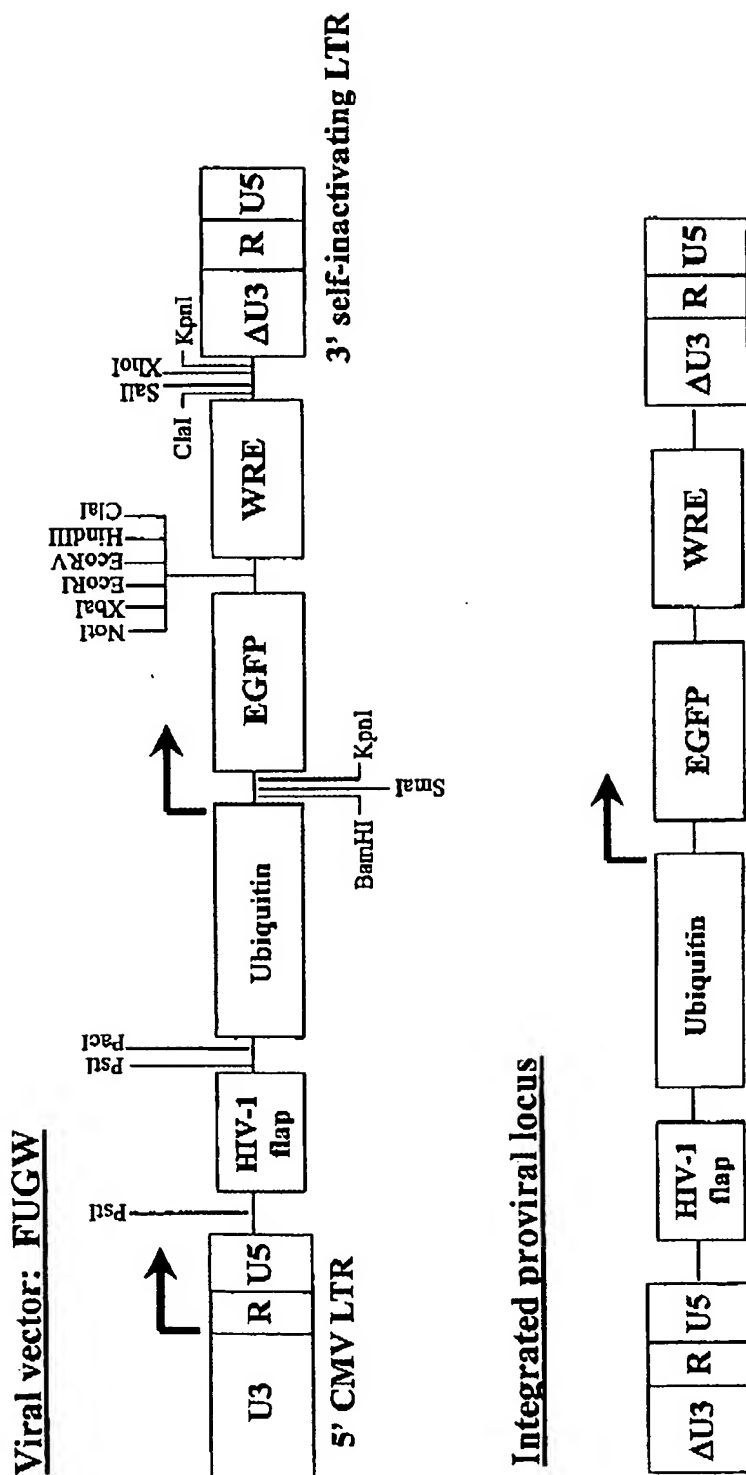
45. A method of identifying genes responsible for a biological process  
20 comprising:

infecting embryonic cells with a modified retrovirus whose genome comprises a self-inactivating lentiviral 3' LTR;

observing a phenotype in the resulting transgenic animal; and

25 identifying a gene that was disrupted by insertion of the provirus by sequencing the nucleic acid flanking the provirus.

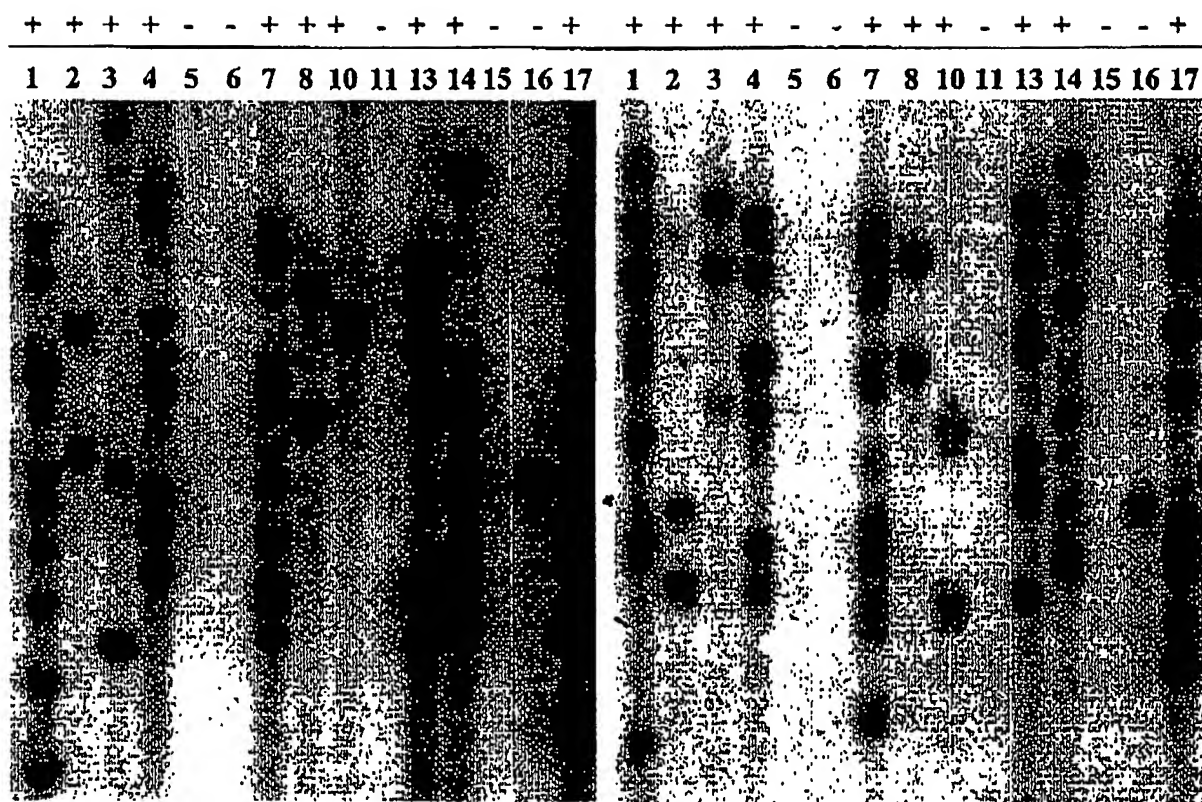
46. The method of claim 45 wherein the genome of the modified retrovirus additionally comprises an origin of replication and an antibiotic resistance gene.

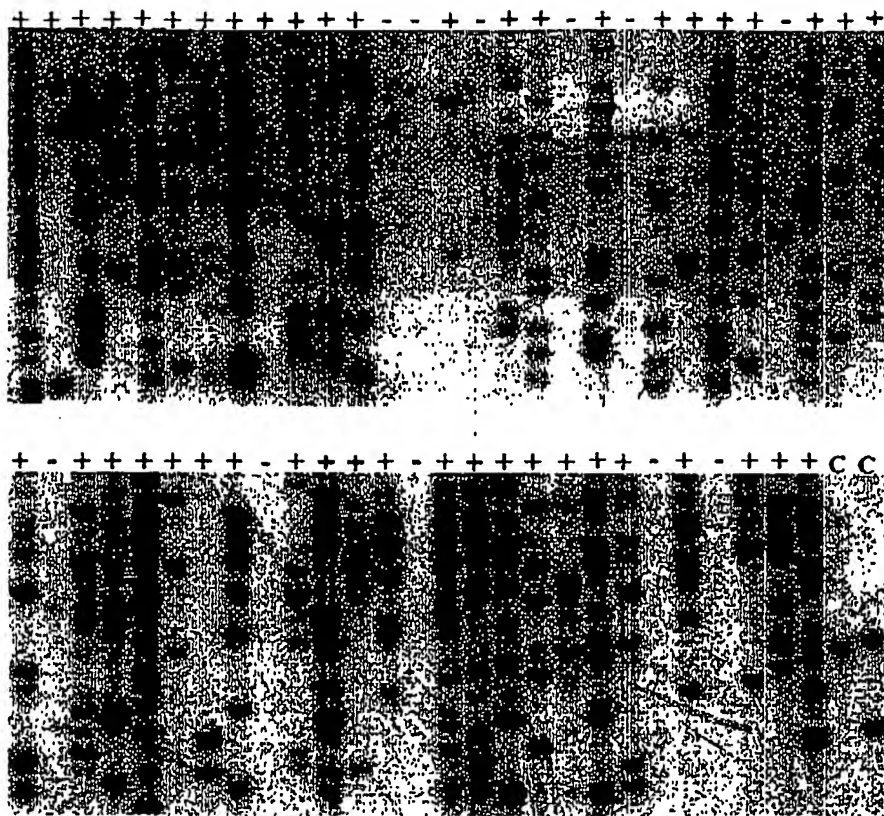


Note: Diagrams are not to scale.

**FIGURE 1**

**FIGURE 2**



**FIGURE 3**

**Ubiquitous GFP expression in rats derived from the delivery of lentiviral vectors to single-cell embryos *in vitro*.** (A) Brightfield (BF) and fluorescent (R3, R5) images of the paws of newborn rats derived from a FUGW-injected embryo. Pup R4, carrying 4 copies of the proviral insert, expresses GFP in the paw, as well as all tissues and organs examined. A littermate (R3) carrying no transgene is included for comparison. (B) Southern blot analysis of proviral insertions in rats generated by injection of FUGW lentivirus into the perivitelline space of single-cell embryos. Genomic DNA was digested with *Pst*I and hybridized with a GFP+WRE probe. Plus signs above each lane indicate GFP expression in the animal detectable by direct viewing under a fluorescent microscope.

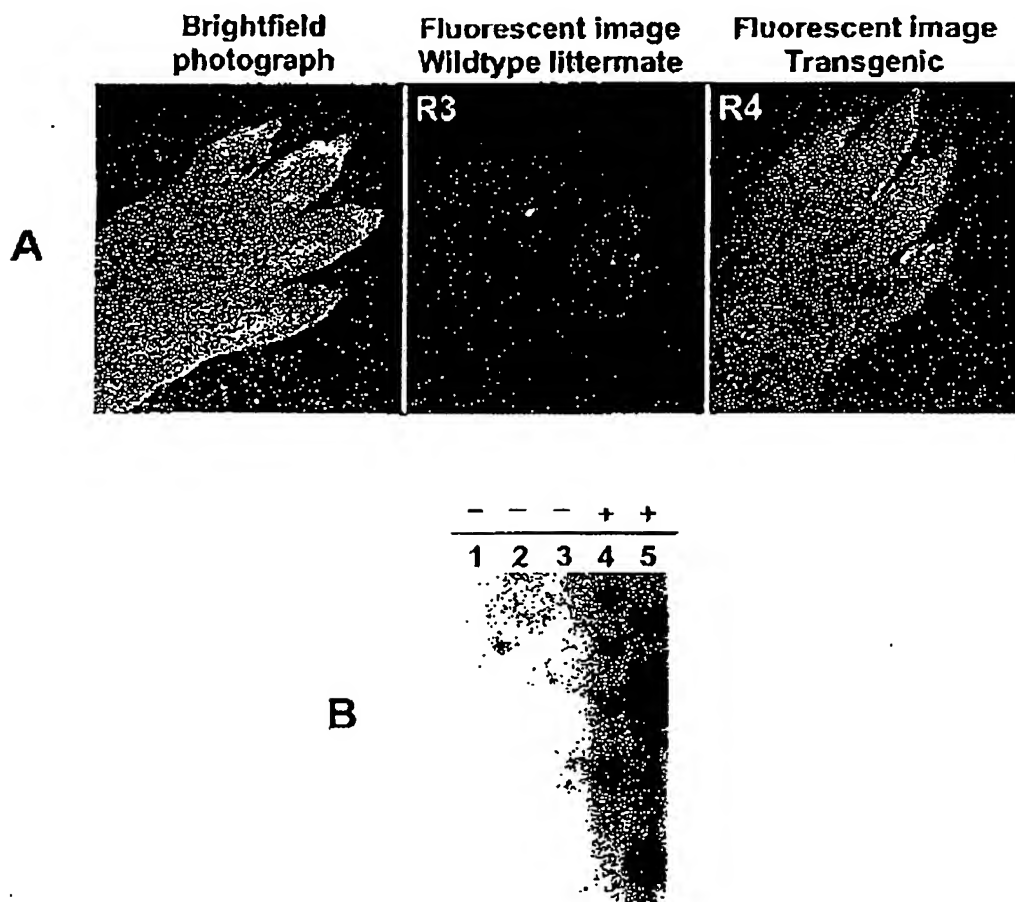
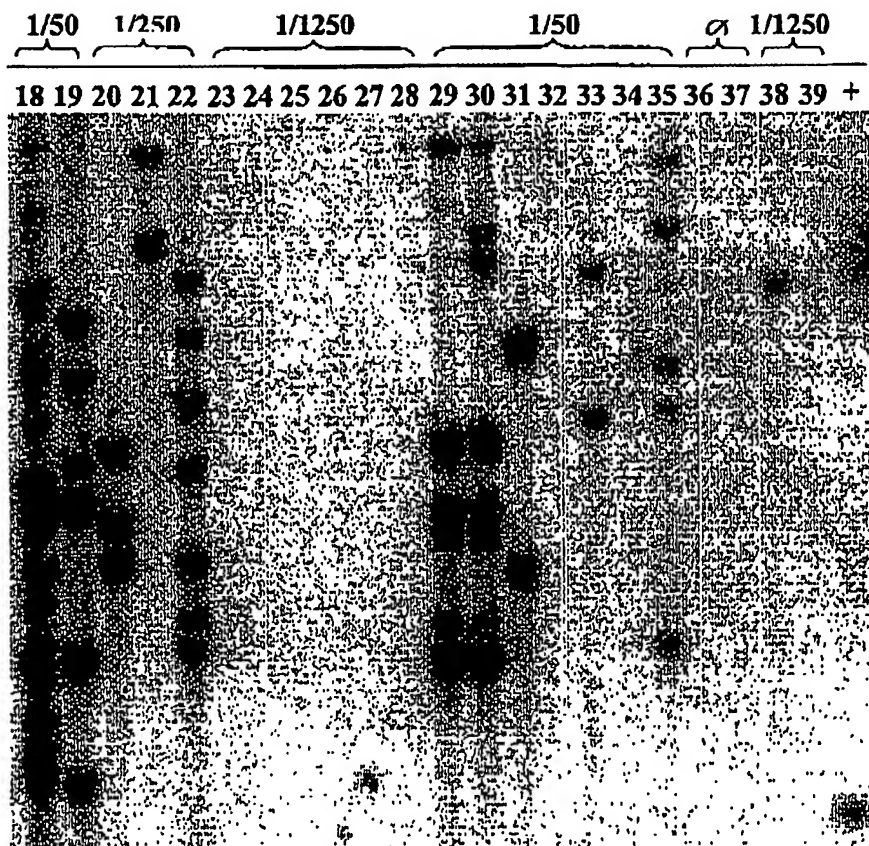


Figure 4

FIGURE 5



**GFP expression in major tissues and organs of a founder mouse.** The animal was perfused intracardially with PBS, pH 7.4, and then 3% paraformaldehyde, and viewed immediately under a Leica fluorescent dissecting microscope. The particular mouse shown here was generated by co-incubation of the denuded embryo with the lentiviral suspension, and carries 8 proviral insertions. A wildtype animal, identically perfused and photographed, is included for comparison.

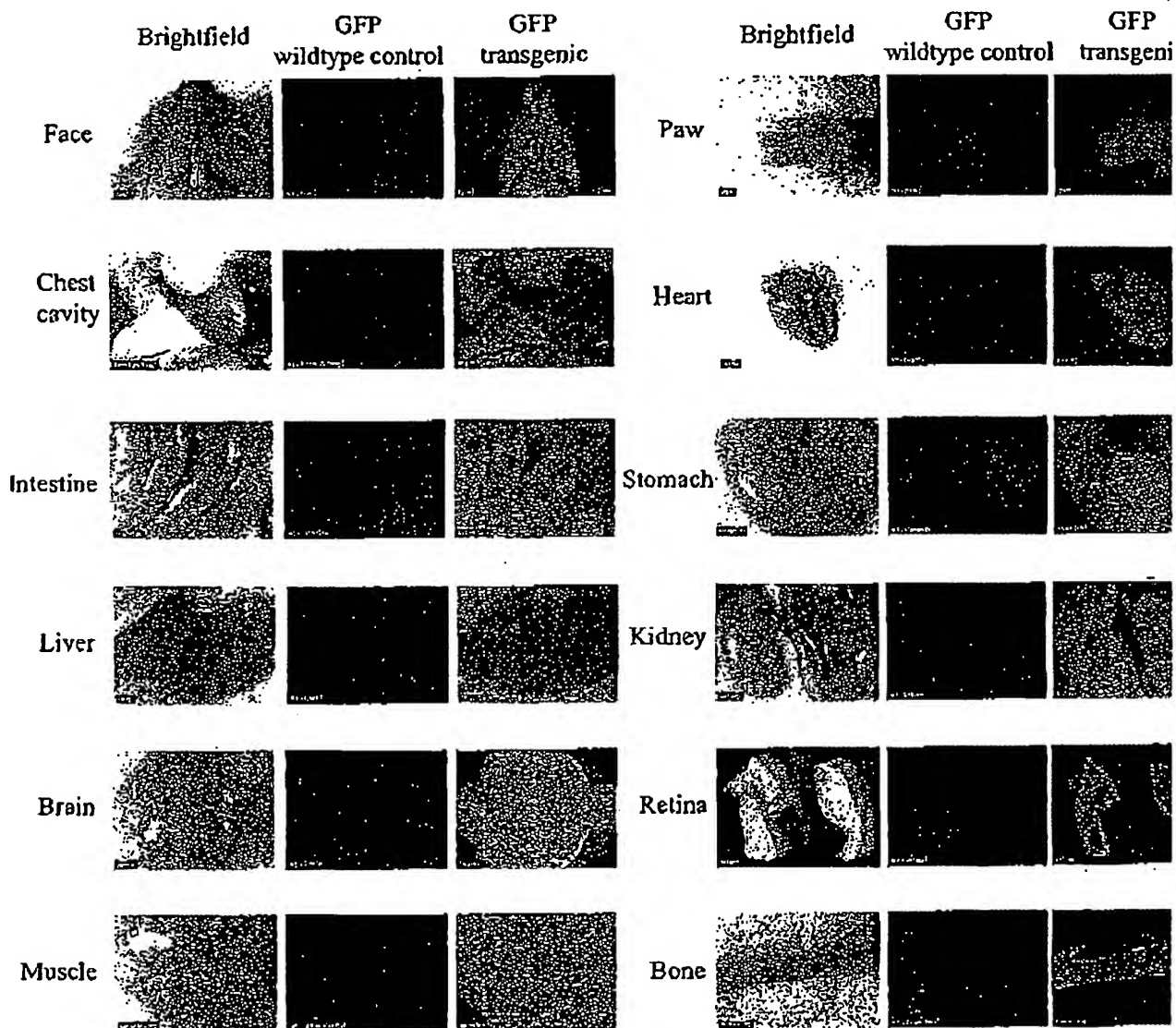


Figure 6





**F1 progeny express GFP.** Transgenic founder mice give rise to transgenic progeny, as detected by GFP expression in the live newborn pup. Transgenesis is also confirmed by Southern analysis. The particular pup imaged here is descended from an animal with 10 proviral insertions.

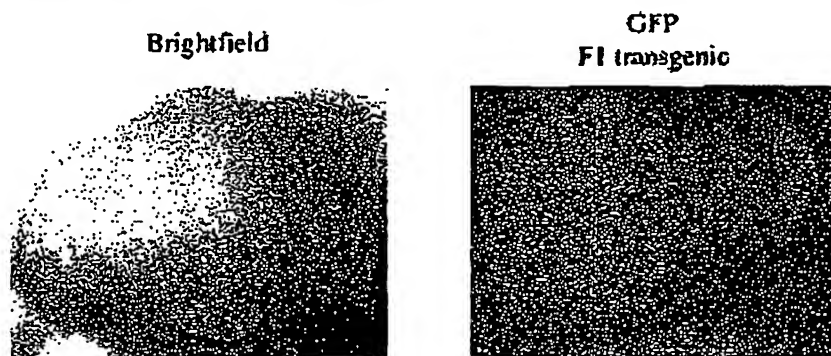
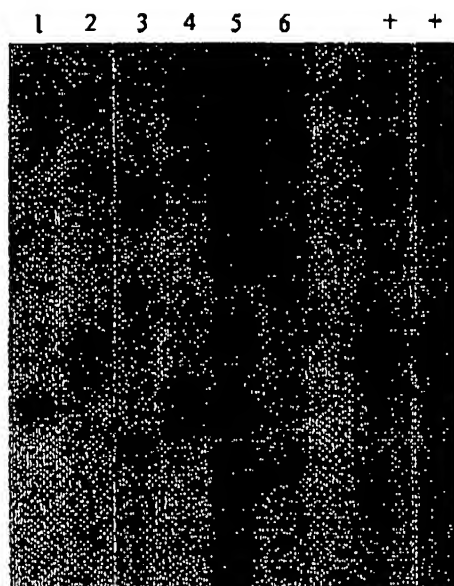


Figure 8

**FIGURE 9**



**GFP expression pattern in an 11.5-dpc mouse embryo derived from the perivitelline space injection of a lentivirus carrying a histone2B-GFP fusion construct under the control of the myogenin promoter (Yee, S.P. and Rigby, P.W.J. (1993). *Genes & Dev.* 7, 1277-89). At this stage in development, myogenin expression has been shown to be limited to those tissues that are committed to the myogenic lineage and give rise to skeletal muscle. GFP expression is localized to the somites and can be seen in the emerging muscles in the limb buds, eye, and jaw.**

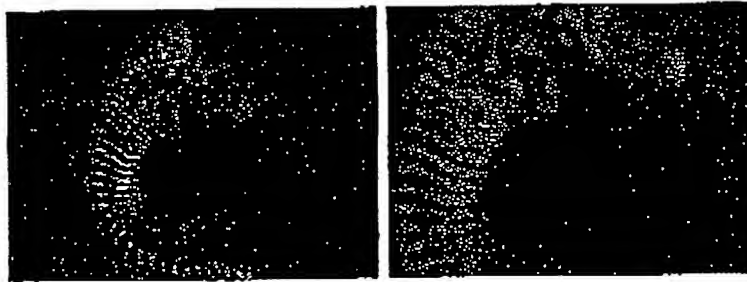
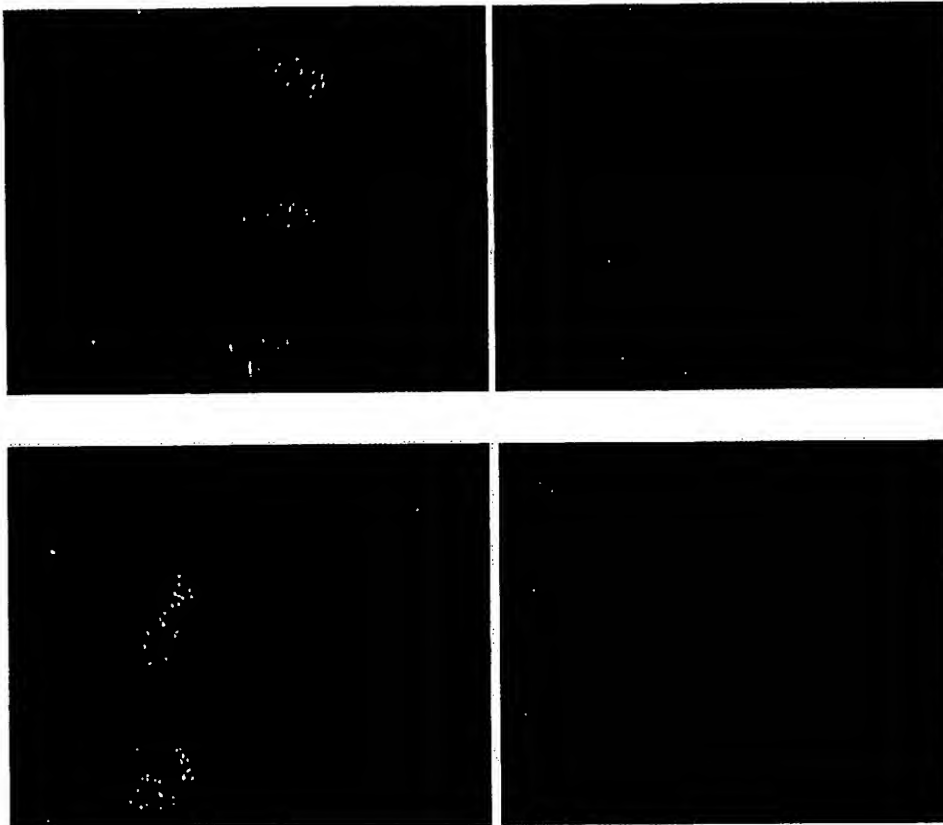


Figure 10

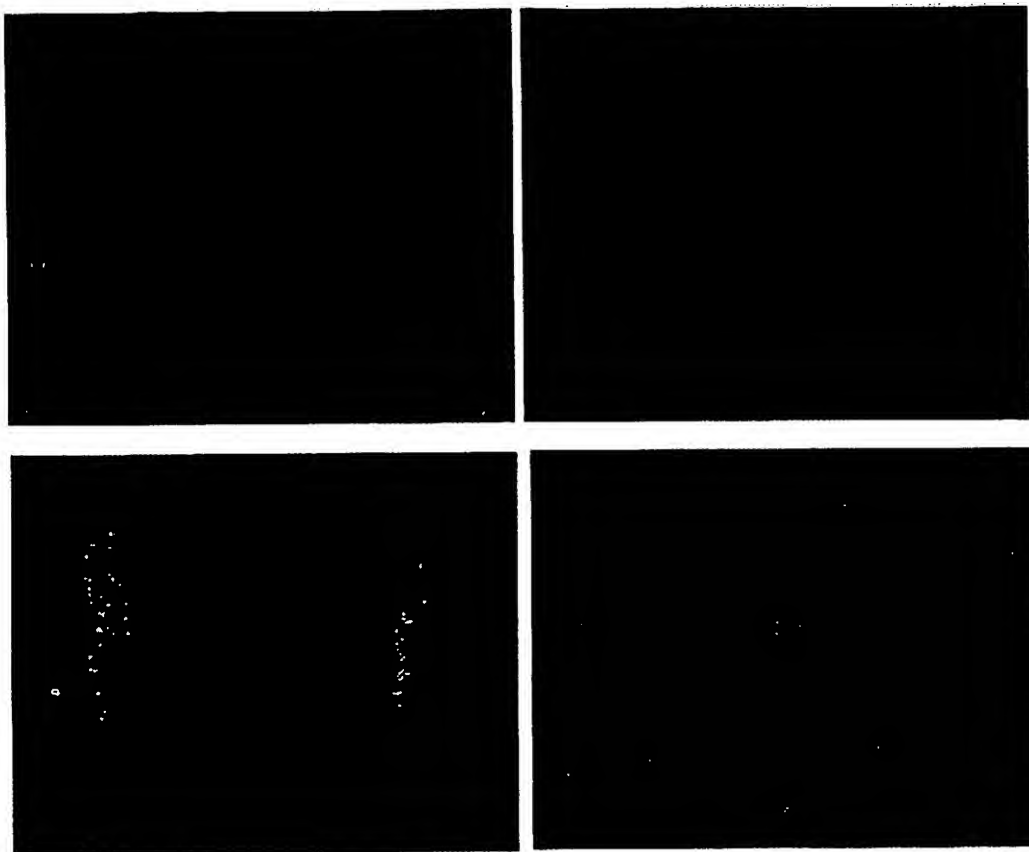
**FIGURE 11**



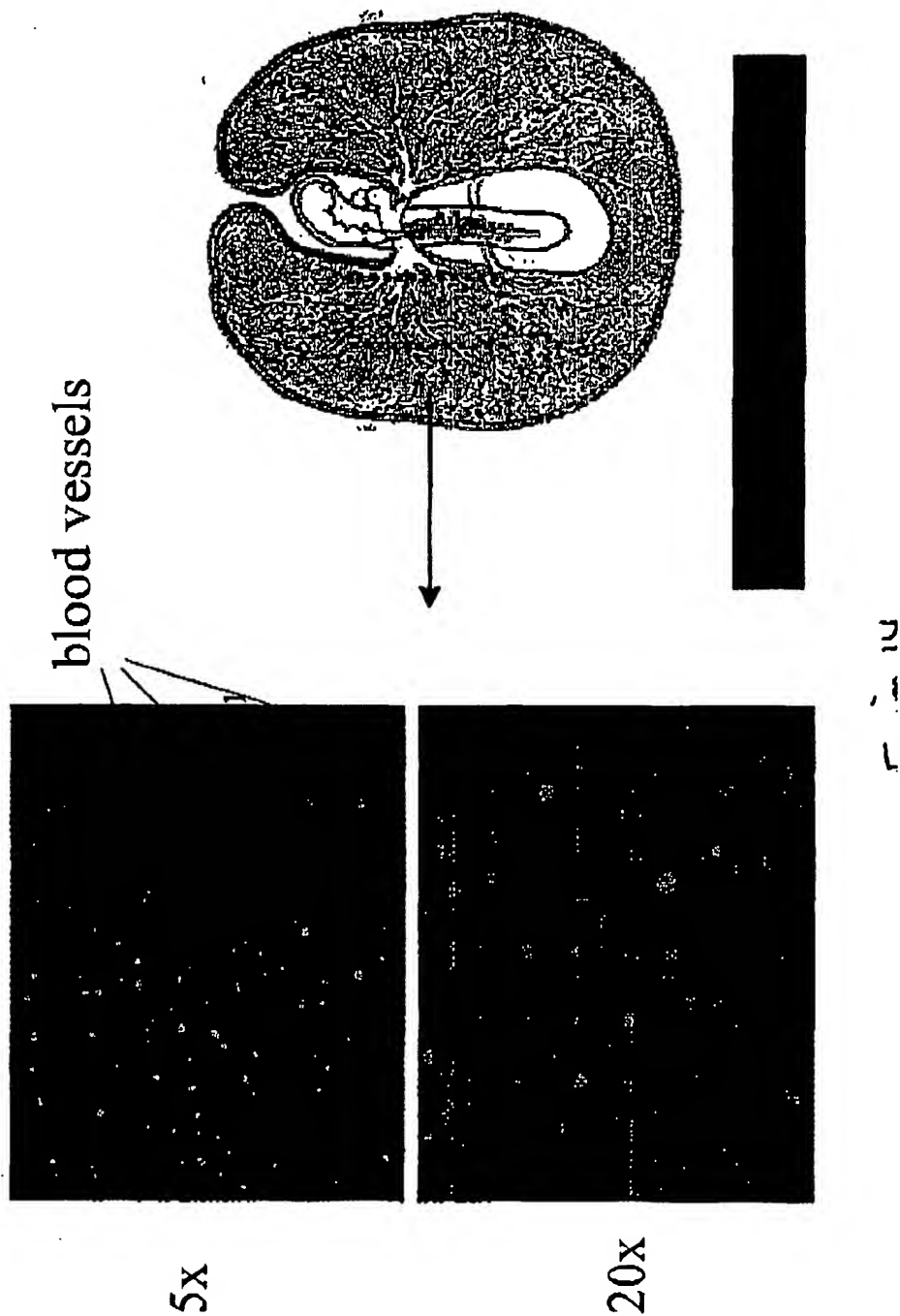
**Figure 12**



**FIGURE 13**

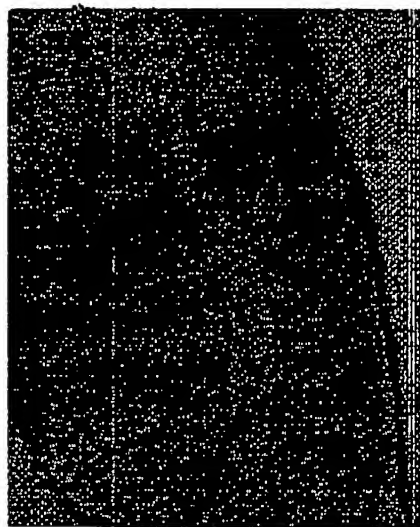
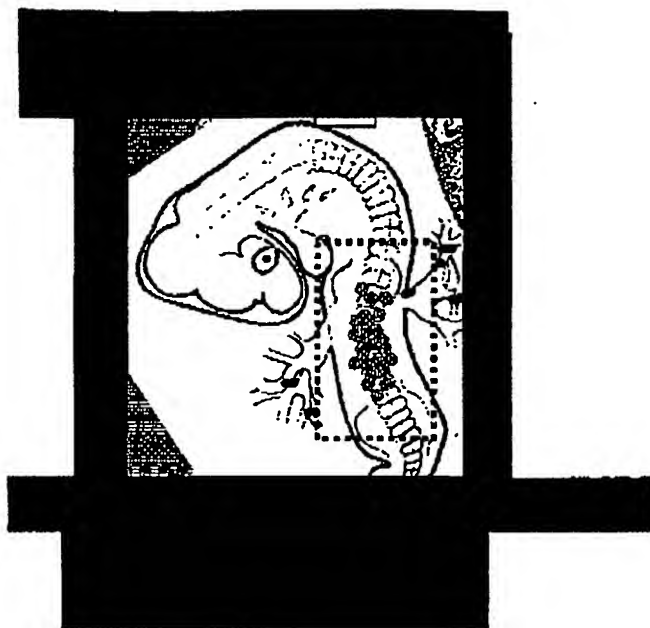


# H2BGFP expression in the extraembryonic tissue

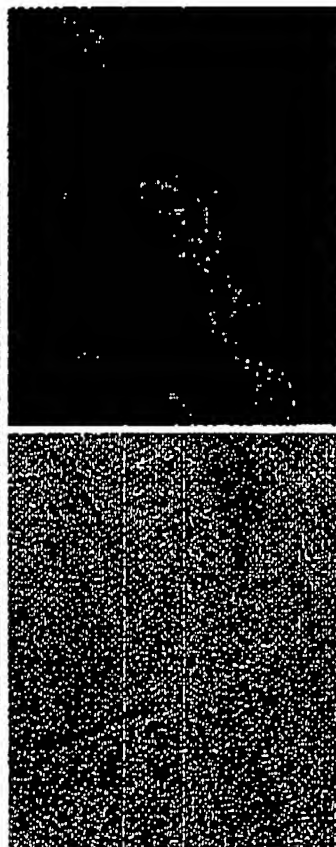




# H2BGFP expression inside of the zebra finch embryo



5x



20x

Figure 15

**Figure 16****GFP**

atggtgagcaaggcgaggagctgttcaccggggtggtgccatcctggtcgagctggacggcgacgtgaacggccacaag  
ttcagcgtgtccggcgaggcgaggcgatgccacctacggcaagctgacctgaagttcatctgcaccaccggcaagctgc  
ccgtgccctggcccaccctcgtgaccaccttcacctacggcgtgcagtgcttcagccgctaccccaccacatgaagcagcac  
gacttctcaagtcgccatgccgaaggctacgtccaggagcgaccatcttctcaaggacgacggcaactacaagaccg  
cgccgaggtgaagttcgaggcgacaccctggtgaaccgcatcgagctgaagggcacgacttcaaggaggacggcaacat  
cctggggcacaagctggagtacaactacaacagccacaacgctctatatcatggccgacaagcagaagaacggcatcaaggtg  
aacttcaagatccgccacaacatcgaggacggcgagcgtgcagctcggcgaccactaccagcagaacacccccatcggcgac  
ggccccgtgctgctgccgacaaccactacctgagcaccagtcggcctgagcaagaccccaacgagaagcgatcac  
atggtcctgctggagttcgtgaccgccggggatcactcacggcatggacgagctgtacaagtaa

**FIGURE 17****H2B-GFP**

accatgccagagccagcgaagtctgctcccgccccgaaaaagggtccaaagaaggcgggtgactaaggcgcagaagaaagg  
cggcaagaagcgcgaagcgcagccgaaggagagctatccatctatgtgtacaagggttctgaagcaggtccacctgacaccg  
gcatttcgtccaaggccatgggcatcatgaattcgtttgtgaacgacatttcgagcgcacgcaggtgaggctcccgctggcg  
cattacaacaagcgtctgaccatcacctccaggagatccagacggccgtgcgctgctgctgctgctggggagttggccaagc  
acgccgtgtccgaggggtactaaggccatcaccaagtacaccagcgttaaggatccaccggcgcaccatggtgagcaaggg  
cgaggagctgttcaccgggggtggtgccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggc  
gaggggcgagggcgatgccacctacggcaagctgacctgaagttcatctgcaccaccggcaagctgccgtgacctggccca  
ccctcgtgaccacctgacctacggcggtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttctcaagtccg  
ccatgcccgaaggctacgtccaggagcgcaccatcttctcaaggacgacggcaactacaagaccgcgccgaggtgaagtt  
cgaggggcgacacctggtgaaccgcatcgagctgaaggggcatcgacttcaaggaggacggcaacatcctggggcacaaagt  
ggagtacaactacaacagccacaacgtctatatacatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgcc  
acaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgc  
ccgacaaccactacctgagcaccagtcgccctgagcaagaccccaacgagaagcgcgatcacatggtcctgctggagtt  
cgtgaccggccggggatcactctcgcatggacgagctgtacaagtaa

**FIGURE 18**HIV NL4.3 flap

acaaatggcagtatccacaaattttaaaagaaaagggggggttggggggtacagtgcaggggaaagaatagtagacataat  
agcaacagacatacaaaactaaagaattacaaaaacaaattacaaaaattcaaaatttcgggtttattacagggacagcagagatc  
cagtttgg

WRE

atcgataatcaacctctggattacaaaatttgtgaaagattgactggattcttaactatgttgctcctttacgctatgttgatacgtg  
ctttaatgcctttgtatcatgctattgcttccgtatggctttcatttctccttctataaatcctgggtgctgtctcttatgaggagttg  
tggcccgttgtaggcaacgtggcgtgggtgtgactgtgttgctgacgcaacccccactgggtggggcattgccaccacgtgc  
agctcctttccgggactttcgtttccccctccctattgccacggcgggaactcatgccgctgccttgcctgctgctggacaggg  
gctcggctgttgggcaactgacaattccgtgggtgtgtcgggggaagctgacgtcctttccatggctgctcgcctgtgttgccacctg  
gattctgcgctgggacgtccttctgtacgtcccttcggccctcaatccagcggaccttccctcccgcggcctgctgccggctctgc  
ggcctctccgcgtcttcgccttcgccctcagacgagtcggatctccctttggggccgcciccccgcctgatcgat

**FIGURE 19**Myogenin promoter

gtctctagctgcatatgtagcagaagatggcctagtcggccatcattgggaagagaggcccttggtattgcaaactatatgccc  
 agtacagggggaacgccagggcccaagaagtgggaatgagtggttaggggagcaggcggggggagggggggttagggaa  
 ctttgggatagcatttgaaatgtaaatgaagaaaatctataaaaaataatttaaaaagagcgtcagacaggggactgaaca  
 gctcttgactaggggagagaaggcaatgtagagtagtctgtgagtctaatccttgctaaacactgacttcacctgacccctacta  
 cttaaggcccccccccttacttaagaagtcctgtgttcttcttacttcaatctaccccccaacatcatgagacctggtcaagaagctg  
 tagaaacccaaaagtgaatccatttgccttctgggttctgtctttgctccatggacgataaggacacacacacacacacacac  
 acacacacacacacacacacgccccaaatctggagtggtcctgatgtgtagtggttaggtctttaggggtctcatgggactgac  
 atagtatggttaaggtgctgctgagcaggaaagagaaggctaaagtggatttcaagaccccttcccgctccgccaagacaaccc  
 ctttctgttcccttctgcccctgtccaccagctgccttggaccatggaggagagagtagggcaggaggcccggttaggagtaatt  
 gaaaggagcagatgagacgggggaatgcaccacccccaccttccctgccccacaggggctgtggagaatgaaaactaat  
 caaattacagccgacggcctcccgacccgtgcacaggagccgctgggcccaggggcaggccctgcagggtgggggtggggg  
 caaaaggagagggaagggaatcacatgaaccactggaaacgtcttgatgtgcagcaacagcttagaggggggctcagggtt  
 ctgtggcgttggctatattatctctgggttcagccagcaggagggtttaaattggcaccagcagttggtgtgaggggctgcgg  
 gagcttgggggcccagtggcaggacaagccttttgcgacctgatg

Lck promoter (partial)

tgagtcaccatgtgattgctgggaattgaactcaagacctctgggaagagcagtcagtgctcttttgtttgtttgtttgtttgg  
 gttttttttgagacaggggttctctgtgtagccgtgtgtgtgggtagccctgtgtgtgagtggtgtgtgtgtgtttgtgtatgttc  
 aaaatagattaaacaactgagagatggaaataggcttcttgacatcaaaaacatgacgtgaacccctttattaaatcaacactcag  
 agacaggagcatctctgcagggttgaggccagcctgctctacagaatgaatttcagggtctcaaggctcagcttggtctacaaagtga  
 gtttcagatctcaaggccagccaagactatgcagtaagaccttagctaaaataaataaataaataaataaataaataaataaagttaatctt  
 cgcttggcaaaccgataattgaggaccagtgctcagggaaggaggcacacgggaattccagaggctacagagggagcctcgtc  
 ctgacctggttagagcaactctactttactggctgtgtctatgagggttctgcttgatttcatttgacaaaaagtccacagctaaacca  
 ggcaaggaggagccgaagtagacacagccacccggggcgcccaacagggttctctctgctgctgagaaagcaaaagcctgtt  
 gaagaaaactctctgaaggagactgtggttagtggtggggtaggggtgctgggggtgggctgaggctgagggtgactctaa  
 ggagctggaaacctctcagcttcggtggctaggcaggggaggttgaatgaaggaggacaggtaccctccttgggtggaggagggt  
 ggaatgaaactctcgtttccccagaacttggcaaaagtgtgtgtgatgtctccaggtagtccccaaaggaggaggctagca  
 gagctggggaggcaggaagtgggtaactagactaacaagatgcctgcctgtggcggttggccatccagggtggaggagggtg  
 ggactagccttgggctgggcctcctgtgaacttgggtgcttgagggtcagagggaaccagtcaggagctgaaatccacgat  
 tcagcgcttctgtctgcggccaatggggcccttgagctgacgatctcgggtacttttgaactccagaacagggtctagggat  
 gctgatgttggggcgagtggttagggccagctccttcaggccctctacattccttcagggatcatgggctgt

**FIGURE 20**Human ubiquitin promoter

gggtgcagcggcctccgcgcgggttttggcgccctcccgcgggcgccccctcctcacggcgagcgctgccacgtcagacg  
aagggcgaggagcgttctgaccttcgcccggacgctcaggacagcggcccgtctcataagactcggccttagaacc  
ccagtatcagcagaaggacatttaggacgggacttgggtgactctagggcactggtttcttccagagagcggaacaggcga  
ggaaaagtagtcccttctcggcgattctcgaggaggatctccgtggggcggtgaacgccgatgattatataaggacgcggg  
gtgtggcacagctagtccgtcgcagccgggatttgggtcgcgggtcttgttggatcgctgtatcgtcacttgggtgagtgcg  
ggctgtcgggtggccggggcttctgtggccgccggggcgctcgggtgggacgggaagcgtgtggagagaccgccaagggt  
gtagtctgggtccgcgagcaagggtgccctgaactgggggttggggggagcgcacaaaatggcggctgttcccagagcttgaa  
tggaagacgcttgtaaggcgggctgtgaggtcgttgaacaagggtggggggcatgttgggcggcaagaaccaaggcttga  
ggccttcgctaatacggggaaagctcttattcgggtgagatgggctggggcaccatctggggaccctgacgtgaagttgtcactg  
actggagaactcgggttgcgtcgtgttgcggggcggcagttatgcgggtgccgttgggcagtgacccgtaccttgggagc  
gcgcgcctcgtcgtgtcgtgacgtcacccgttctgttggcttataatgcaggggtggggccacctgccggtagggtgcggtagg  
cttttctccgtcgcaggacgcagggttcgggcctagggttaggtctcctgaatcgcagggcggggacctctggtgaggggag  
ggataagtgaggcgtcagtttcttgggtcgttttatgtacctatcttctaagtagctgaagctccggtttgaactatgcgctcggg  
gttggcgagtggtttgtgaagtttttaggcaccttttgaatgtaatcatttgggtcaatatgtaatttcagtgtagactagtaaatt  
gtccgctaaattctggccgttttggcctttttgtagaca

ctcagacaaatggcaglattatccacaattttaaaagaaaagggggattgggggtacagtgcaggggaaagaatagtag  
acataatagcaacagacatacaaaactaaagaattacaaaaacaaattacaaaatttcgggtttattacagggacagc  
agagatccagtttgctgcagtttaataaagaatcgtgggtgcagcggccctccgcgcgggggtttggcgcctccgcgcggcgcccc  
cctcctcacggcgagcgctgccacgtcagacgaaagggcgcaaggagcgttctgatactctccgccggagcgtcaggacagc  
ggcccgcgtctcataagactcggccttagaaccacgatatcagcagaaggacatttagacgggacttgggtgactctagggc  
actggttttcttcagagagcgggaacaggcgagggaaaagtatgcccttctcggcgattctgcggagggatctccgtggggcg  
tgaacgccgatgattatataaggacgcgcgggggtgtggcacagctagtctcgtcgcagccgggattgggtcgcgggtctctgtt  
gtggatcgctgtgatcgtcacttggtagttgcgggctgtctgggtcggccggggcttctgtggccggcgggcggtcgggtggg  
acgggaagcgtgtggagagaccgccaagggtctgtagtctgggtccgcgagcaagggttgcctgaactgggggttgggggag  
cgacaaaatggcggtgttcccgagtcctgaatggaagacgcttgaaggcgggcgtgtgaggtcgttgaacaagggtggggg  
gcatgtgtgggcggaagaacccaaggctctgaggccttcgtaatgcgggaaagctcttattcgggtgagatgggctggggca  
ccatctgggggacctgacgtgaagttgtcactgactggagaactcgggttctgcgtcgttgggtcggggcgggcagttatcggt  
ggcgttgggcagtgacccgtacctttgggagcgcgcgcctcgtcgtgtcgtgacgtcacccgttctgttggctataatgcagg  
gtggggccacctgccggtaggtgtgcggtaggcttttccgtcgcaggacgcagggttcgggacctagggttaggctctctgaa  
tcgacaggcgccggaccctgtgtgaggggagggataagtgaggcgtcagtttcttgggttattgtacctatctcttaagta  
gtgaaagctccgggtttgaactatgcgctcggggttggcgagtgtgtttgtgaagtttttaggcacctttgaaatgtaatcatttg  
gtcaatatgtaatttcagtgtagactagtaaattgtccgctaattctggccgtttttggctttttgttagacaaagcttctgcaggtc  
gactctagaggatccccgggggtaccatggtgagcaaggcgaggagcgtgttaccgggggtgtgtgccatctgtgtcgaagc  
tggacggcgacgtgaacggccacaagttcagcgtgtccggcgagggcgaggcgatgccacctacggcaagctgacctga  
agttcatctgcaccaccggcaagctgccgtgccctggcccacctcgtgaccacctcacctacggcgtgagtgcttcagcc  
gtaccccgaccacatgaagcagcacgactcttcaagtccgcctatgccggaaggctacgtccaggagcgacacatcttctca  
aggacgacggcaactacaagaccgcgcggaggtgaagtgcaggggcgacacctgtgtgaaccgcatcgagctgaagggc  
atcgacttcaaggagggacggcaacatctggggcacaagcttgagtacaactacaacagccacaacgtctatatcatggccga  
caagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgcggaccacta  
ccagcagaacacccccatcggcgacggccccgtgctgtgtcccgacaaccactacctgagcaccagctcggcctgagcaa  
agaccccaacgagaagcgcgatcatggtctgtcgtgaggttcgtgaccggccggggtacacacagggcatgagcaggtg  
tacaagtaagcgggcgtctagagaattcgatcaagcttaicgatacgaatacaacctctggattacaaaattgtgaaagatt  
gactggtattcttaactatgtgtccttttacgctatgtggatagcgtcttaatgcctttgtatcatgtattgttccggtatggctt  
cattttctctccttgtataaactcgtggtgtgtctctttatgaggagttgtggccgttgtcaggcaacgttggcgttgggtgtcactgt  
gtttgtgtacgcaacccccactgggttggggcattgccaccacctgtcagctcctttccgggacttgccttccccctccctattgcc  
acggcggaactcatcggccgtccttgcccgctgctggacagggggctcggctgttgggcactgacaattccgtgtgtgtgtc  
gggaaagctgacgtcctttccatggctgtctgcctgtgttgcacctggattctgcgcgggacgtccttctgtactgtccttcggc  
ctcaatccagcggaccttcttcccgcggtgtgtcgggtctgtcggccttctccgctcttcgccttcgccttcagacgagtc  
ggatctcccttggggcgctccccgcctgaicgatalcgtatgtcgaacctgaggggtacc

## SEQUENCE LISTING

<110> California Institute of Technology  
BALTIMORE, DAVID  
HONG, ELIZABETH J.  
LOIS-CABALLE, CARLOS  
PEASE, SHIRLEY

<120> METHOD FOR PRODUCING TRANSGENIC ANIMALS

<130> CALTE.006VPC

<150> 60/322,031

<151> 2001-09-13

<150> 60/347,782

<151> 2002-01-09

<160> 8

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 178

<212> DNA

<213> Human immunodeficiency virus

<400> 1

```
acaaatggca gtattcatcc acaattttta aagaaaaggg gggattgggg ggtacagtgc 60
aggggaaaga atagtagaca taatagcaac agacatacaa actaaagaat tacaaaaaca 120
aattacaaaa attcaaaatt ttcgggttta ttacaggggac agcagagatc cagtttgg 178
```

<210> 2

<211> 1221

<212> DNA

<213> Homo sapiens

<400> 2

```
gggtgcagcg gcctccgcgc cgggttttgg cgctccgcgc gggcgcccc ctcctcacgg 60
cgagcgctgc cacgtcagac gaaggcgca ggagcggtcc tgatccttcc gcccgacgc 120
tcaggacagc ggcccgtgc tcataagact cggccttaga accccagtat cagcagaagg 180
acattttagg acgggacttg ggtgactcta gggcactggg tttctttcca gagagcggaa 240
caggcgagga aaagtagtcc cttctcggcg attctgcgga gggatctccg tggggcgggtg 300
aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagtcccg tcgcagccgg 360
gatttgggtc gcggttcttg tttgtggatc gctgtgatcg tcacttggtg agttgcgggc 420
tgctgggctg gccggggctt tcgtggccgc cgggccgctc ggtgggacgg aagcgtgtgg 480
agagaccgcc aagggtgta gtctgggtcc gcgagcaagg ttgccctgaa ctggggggtg 540
gggggagcgc acaaaatggc ggctgttccc gactcttgaa tggaaagacg ttgtaaggcg 600
ggctgtgagg tcgttgaaac aagggtgggg gcatgggtgg cggcaagaac ccaaggctct 660
gaggccttcg ctaatgcggg aaagctctta ttcgggtgag atgggctggg gcaccatctg 720
gggaccctga cgtgaagttt gtcactgact ggagaactcg ggtttgtcgt ctggttgccg 780
gggcggcagt tatgcggtgc cgttgggcag tgcaccgta cctttgggag cgcgcgcctc 840
gtcgtgtcgt gacgtcaccg gttctgttgg cttataatgc aggggtggggc cacctgccgg 900
taggtgtcgt gtaggctttt ctccgtcgca ggacgcaggg ttcgggccta gggtaggctc 960
tcctgaatcg acaggcgccg gacctctggt gaggggaggg ataagtgagg cgtcagtttc 1020
```



```

tttggctcggg tttatgtacc tatcttctta agtagctgaa gctccggttt tgaactatgc 1080
gctcgggggtt ggcgagtgtg ttttgtgaag ttttttaggc accttttgaa atgtaatcat 1140
ttgggtcaat atgtaatttt cagtgttaga ctagtaaatt gtccgctaaa ttctggccgt 1200
ttttggcttt tttgttagac a                                     1221

```

<210> 3  
 <211> 720  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> This sequence encodes for a green fluorescent protein variant.

```

<400> 3
atggtgagca agggcgagga gctgttcacc ggggtggtgc ccacccctgt cgagctggac 60
ggcgacgtga acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccggtgcc ctggcccacc 180
ctcgtgacca ccttcaccta cggcgtgcag tgcttcagcc gctaccccca ccacatgaag 240
cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagaccgcg gccgaggtga agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagca cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggatc actcacggca tggacgagct gtacaagtaa 720

```

<210> 4  
 <211> 604  
 <212> DNA  
 <213> Woodchuck hepatitis virus

```

<400> 4
atcgataatc aaactctgga ttacaaaatt tgtgaaagat tgactggtat tcttaactat 60
gttgctcctt ttacgctatg tggatacgtc gctttaatgc ctttgtatca tgctattgct 120
tcccgatagg ctttcatttt ctccctcttg tataaatcct gggtgctgtc tctttatgag 180
gagttgtggc cgtttgtcag gcaacgtggc gtggtgtgca ctgtgtttgc tgacgcaacc 240
cccactggtt ggggcattgc caccacctgt cagctccttt ccgggacttt cgctttcccc 300
ctccctattg ccacggcgga actcatcgcc gcctgccttg ccgctgctg gacaggggct 360
cggctgttgg gcaactgaaa ttccgtgggt ttgtcgggga agctgacgtc ctttccatgg 420
ctgctcgctt gtgttgccac ctggattctg cgcgggacgt ccttctgcta cgctcccttcg 480
gccctcaatc cagcggacct tccttcccgc ggctgctgct cggtctgctg gcctcttccg 540
cgtcttcgcc ttgcctctca gacgagtcgg atctcccttt gggccgcctc cccgcctgat 600
cgat                                     604

```

<210> 5  
 <211> 1119  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> This sequence encodes for a fusion protein between human histone 2B protein and a green fluorescent protein variant.

&lt;400&gt; 5

```

accatgccag agccagcgaa gtctgctccc gccccgaaa agggctccaa gaaggcggtg 60
actaaggcgc agaagaaagg cggcaagaag cgcaagcgca gccgcaaggga gagctattcc 120
atctatgtgt acaaggttct gaagcaggtc caccctgaca ccggcatttc gtccaaggcc 180
atgggcatca tgaattcgtt tgtgaacgac attttcgagc gcatcgagcag tgaggcttcc 240
cgcttggggc attacaacaa gcgctcgacc atcacctcca gggagatcca gacggccgtg 300
cgctgtgtgc tgcctgggga gttggccaag cacgccgtgt ccgagggtac taaggccatc 360
accaagtaca ccagcgctaa ggatccaccg gtcgccacca tggtaggcaa gggcgaggag 420
ctgttcaccg ggggtgtgcc catcctggtc gagctggacg ggcacgtaaa cggccacaag 480
ttcagcgtgt ccggcgaggg cgaggcgat gccacctacg gcaagctgac cctgaagttc 540
atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac cctgacctac 600
ggcgtgcagt gcttcagcgc ctaccccgac cacatgaagc agcacgactt cttcaagtcc 660
gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga cggcaactac 720
aagaccgcgc ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat cgagctgaag 780
ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta caactacaac 840
agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt gaacttcaag 900
atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca gcagaacacc 960
cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac ccagtccgcc 1020
ctgagcaaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc 1080
gccgggatca ctctcgcat ggacgagctg tacaagtaa 1119

```

&lt;210&gt; 6

&lt;211&gt; 1276

&lt;212&gt; DNA

&lt;213&gt; Murine

&lt;400&gt; 6

```

tgagtaccca tgtgattgct gggaattgaa ctcaagacct ctggaagagc agtcagtgtc 60
ctttttgttt gtttgtttgt ttgtttgttt ggggtttttt tttgagacag ggtttctctg 120
tgtagccgtg tgtgtggggt agccctgtgt gtgagtgtgt gtgtgtgtgt ttgtggtatg 180
ttgcaaaata gattaaacaa ctgagagatg gaataggtct tcttgacatc aaaaacatga 240
tcgtgaaccc ctttattaaa tctaactctc agagacagga gcatctctgc aggtttgagg 300
ccagcctgct ctacagaatg aatttcaggt ctcaaggtca gcttggtcta caaagtgagt 360
ttcagatctc aaggccagcc aagactatgc agtaagacct tagctaaaat aaataaataa 420
ataaaataaa ataaaagtta atcttcgctt ggcaaaccga taattgagga ccagtgtca 480
ggaaggaggc acacgggaat tccagaggct acagaggag cctcgctctg acctgggttag 540
agcaactcta ctttactggc tgtgtctatg aggttctgct tgatttcatt tgacaaaaag 600
tttccacagc taaaccaggc aaggagccg aagtagacac agccaccggg gccgcgcca 660
acaggtttct ctctgctgct gagaagcaaa agcctgtttg aagaaactct ctgaaggaga 720
ctgtggttga gtggtggggg taggggtgct ggggttggc tgaggctgag ggttgactct 780
aaggagctgg aacctctcag cttcgggtgg taggcagggg agttgtaatg aagagggaca 840
ggtacctcc ttggtggagg aggggtggaat gaaactctc gtttcccca gaacttgga 900
aagtgtgtgt gatgtctccc aggtagtccc ccaaaggagg aggctagcag agctggggag 960
gcaggaagtg ggtaactaga ctaacaaaga tgccctgctg tggcggtttg cccatcccag 1020
gtgggagggg gggactagcc ttgggcctgg gcctcctgtg aacttggtgc ttgagggtc 1080
agagggaacc cagtcaggag cttgaatccc acgattcagc gcttctgtct gcggccaatg 1140
ggggcctctg agctgacgat ctcggtact ttttgtaact tccagaacag ggctctagga 1200
tgtctgatgt tggggcgagt ggcttagggc cagctccttc aggcctctct acattccttc 1260
agggatcatg ggctgt 1276

```

&lt;210&gt; 7

&lt;211&gt; 1140

&lt;212&gt; DNA

&lt;213&gt; Murine

&lt;400&gt; 7

```

gtctctagct gcatatgtag cagaagatgg cctagtcggc catcattggg aagagaggcc 60

```

```

ccttggtatt gcaaactata tgccccagta caggggaacg ccagggccaa gaagtgggaa 120
tgagtgggta ggggagcagg gcggggggag ggggggtag ggaacttttg ggatagcatt 180
tgaaatgtaa atgaagaaaa tatctaataa aaaataattt aaaaaagagc gtcagacagg 240
ggactgaaca gctcttgact aggggagaag aaggcaatgt agagtagtct gtgagttcta 300
atccttgcta aacactgact tcacctgacc cctactactt aaggccccc cccttactta 360
agaagtcctt gtgttctctt acttcaatct accccaaca tcatgagacc tggtaaaga 420
agctgtagaa acccaaaagt tgaatccatt tgccctctg ggtttctgtc tttgcctcca 480
tggacgatag ggacacacac acacacacac acacacacac acacacacac acgccccaaa 540
tctggagtgg tctgatgtg gtagtgtag gtctttagg gtctcatggg actgacatag 600
tatggtttaa ggtgctgctg agcaggaaag agaaggctaa gtggattttc aagaccctt 660
cccgtccgtc caagacaacc cctttcttgt tcccttctg ccctgtccac cagctgcctt 720
ggaccatgga ggagagagta ggcaggagc cgggtagga gtaattgaaa ggagcagatg 780
agacggggga atgcaccac cccaccttc cctgcccac aggggctgtg gagaaatgaa 840
aactaatcaa attacagccg acggcctccc gaccgtgca caggagccg ctgggccagg 900
ggcagcctg cagggtggg tgggggcaaa aggagaggga aggggaatca catgtaacca 960
ctggaacgt cttgatgtg agcaacagct tagaggggg ctcaggtttc tgtggcgtt 1020
gctatattta tctctgggtt catgccagca gggagggtt aaatggcacc cagcagttg 1080
tgtgaggggc tgcgggagct tgggggcccag tggcaggaa aagccttttg cgacctgatg 1140

```

<210> 8

<211> 2853

<212> DNA

<213> Artificial Sequence

<220>

<223> This sequence represents a construct comprising a human immunodeficiency virus sequence, a green fluorescent protein variant sequence, a human ubiquitin promoter sequence and a woodchuck hepatitis regulator element sequence.

<400> 8

```

ctgcagacaa atggcagtat tcatccacaa ttttaaaaga aaagggggga ttggggggta 60
cagtgcaggg gaaagaatag tagacataat agcaacagac atacaaacta aagaattaca 120
aaaacaaatt acaaaaattc aaaattttcg ggtttattac agggacagca gagatccagt 180
ttggctgcag ttaattaaag atctgggtgc agcgccctcc gcgcggggtt ttggcgccctc 240
ccgcggggcg cccctctctc acggcgagcg ctgccacgtc agacgaaggg cgcaggagcg 300
ttcctgatcc ttccgcccgg acgctcagga cagcgcccgc ctgctcataa gactcggcct 360
tagaacccca gtatcagcag aaggacattt taggacggga cttgggtgac tctagggcac 420
tggttttctt tccagagagc ggaacaggcg aggaaaagta gtcccttctc ggcgattctg 480
cggagggatc tccgtggggc ggtgaacgcc gatgattata taaggacgcg ccgggtgtgg 540
cacagctagt tccgtcgag ccgggatttg ggtcgcggtt cttgtttgtg gatcgctgtg 600
atcgtcactt ggtgagttgc gggctgctgg gctggccggg gctttcgtgg ccgccgggcc 660
gctcgggtgg acggaagcgt gtggagagac cgccaagggc tgtagtctgg gtccgcgagc 720
aaggttgccc tgaactgggg gttgggggga gcgcacaaaa tggcggctgt tcccagatct 780
tgaatggaag acgcttgtaa ggcggtgtg gaggtcgttg aaacaagggt gggggcatgg 840
tgggcggcaa gaaccgaagg tcttgaggcc ttcgctaatt cgggaaagct cttattcggg 900
tgagatgggc tggggcacca tctggggacc ctgacgtgaa gtttgtcact gactggagaa 960
ctcgggtttg tggtctggtt gcggggggcg cagttatgag gtgccgttg gcagtgacac 1020
cgtacctttg ccgacgcgcg cctcgctcgt tgcgtacgtc acccgttctg ttggcttata 1080
atgcagggtg gggccacctg ccggtagggt tgcggtaggc ttttctccgt cgcaggacgc 1140
agggttcggg cctagggtag gctctcctga atcgacagg gccggacctc tggtagggg 1200
agggataagt gaggcgtcag tttcttttgt cggttttatg tacctatctt ctttaagtagc 1260
tgaagctccg gttttgaact atgcgctcgg ggttggcgag tgtgttttgt gaagttttt 1320
aggcaccttt tgaaatgtaa tcatttgggt caatatgtaa ttttcagtgt tagactagta 1380
aattgtccgc taaattctgg ccgttttttg cttttttgtt agacaaagct tctgcaggtc 1440

```

```

gactctagag gatcccccg ggggtaccatg gtgagcaagg gcgaggagct gttcaccggg 1500
gtggtgcccc tcctggtcga gctggacggc gacgtgaacg gccacaagtt cagcgtgtcc 1560
ggcgagggcg agggcgatgc cacctacggc aagctgaccc tgaagttcat ctgcaccacc 1620
ggcaagctgc ccgtgccctg gccaccctc gtgaccacct tcacctacgg cgtgcagtgc 1680
ttcagccgct accccgacca catgaagcag cacgacttct tcaagtccgc catgcccga 1740
ggctacgtcc aggagcgac catcttcttc aaggacgacg gcaactacaa gaccgcgcc 1800
gaggtgaagt tcgagggcga caccctggtg aaccgcatcg agctgaaggg catcgacttc 1860
aaggaggacg gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc 1920
tatatcatgg ccgacaagca gaagaacggc atcaagggtga acttcaagat ccgccacaac 1980
atcgaggacg gcagcgtgca gctcgccgac cactaccagc agaacacccc catcggcgac 2040
ggccccgtgc tgctgccga caaccactac ctgagcacc agtccgccct gagcaaagac 2100
cccaacgaga agcgcgatca catggtcctg ctggagtctg tgaccgccgc cgggatcact 2160
cacggcatgg acgagctgta caagtaagcg gccgctctag agaattcgat atcaagctta 2220
tcgatatcga taatcaacct ctggattaca aaatttgtga aagattgact ggtattctta 2280
actatgttgc tccttttacg ctatgtggat acgtgcttt aatgcctttg tatcatgcta 2340
ttgcttcccg tatggcttcc attttctcct ccttgataaa atcctggttg ctgtctcttt 2400
atgaggagtt gtggcccggt gtcaggcaac gtggcggtgt gtgcactgtg tttgctgacg 2460
caacccccac tgggtggggc attgccacca cctgtcagct cctttccggg actttcgctt 2520
tccccctccc tattgccacg gcggaactca tcgccgcctg ccttgcccgc tgctggacag 2580
gggctcggct gttgggcact gacaattccg tgggtgtgtc ggggaagctg acgtcctttc 2640
catggctgct cgctgtgtt gccacctgga ttctgcgcgg gacgtccttc tgctacgtcc 2700
cttcggccct caatccagcg gaccttctt cccgcggcct gctgccggct ctgcggcctc 2760
ttcgcgtct tcgccttcgc cctcagacga gtcggatctc cctttgggcc gcctccccgc 2820
ctgatcgata tcgatgtcga cctcgagggt acc 2853

```